

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 85.4096 Seconds
(without alignments)
12875.302 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 639

Sequence: 1 gctgtagcgaggggctggg.....attctctcaggttttaaaa 1946

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool_P/US09270437/runat_06082004_141220_1795/app_query.fasta_1.5582
-DB=A.Geneseq.29Jan04 -QFMT=fastan -SUFFIX=oligo806.rag -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09270437@cgn 1 185 -runat_06082004_141220_1795
-NCFU=6 -ICFU=3 -NO_WMAP -LARGEQJURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOP -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq.29Jan04.*

1: geneseqp1980s.*
2: geneseqp1980s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	216	33.8	261	AAU16161	AAU16161 Human nov
2	216	33.8	261	ABU55230	ABU55230 Human nov
3	216	33.8	319	AAAM3826	AAAM3826 Human pol
4	166	26.0	577	AAAY30649	AAAY30649 A murine
5	116	18.2	250	AAU16579	AAU16579 Human nov
6	116	18.2	250	ABU55648	ABU55648 Human nov
7	36	5.6	579	AAAB11328	AAAB11328 Human lun
8	36	5.6	579	AAAB11365	AAAB11365 Human lun
9	36	5.6	579	ABB74960	ABB74960 Human lun
10	36	5.6	579	ABB75053	ABB75053 Human lun

11	36	5.6	579	5	ABB74997	ABB74997 Human lun
12	36	5.6	579	5	ABB75054	ABB75054 Human lun
13	36	5.6	579	5	ABP61917	ABP61917 Human lun
14	36	5.6	579	5	ABP61974	ABP61974 Human lun
15	36	5.6	579	5	ABP61880	ABP61880 Human lun
16	36	5.6	579	5	ABP61973	ABP61973 Human lun
17	36	5.6	579	7	ADA28536	ADA28536 Recombina
18	36	5.6	579	7	ADA28539	ADA28539 Recombina
19	36	5.6	579	7	ADA28438	ADA28438 Human lun
20	36	5.6	579	7	ADA28266	ADA28266 Human lun
21	36	5.6	579	7	ADD14066	ADD14066 Human src
22	36	5.6	579	7	ADE53471	ADE53471 Human lun
23	36	5.6	579	4	ABG12592	ABG12592 Novel hum
24	36	5.6	586	5	ABP75048	ABP75048 Human lun
25	36	5.6	586	5	ABP61968	ABP61968 Human lun
26	36	5.6	586	7	ADA28517	ADA28517 Recombina
27	36	5.6	619	4	ABG21963	ABG21963 Novel hum
28	24	3.8	148	4	ABG21962	ABG21962 Novel hum
29	21	3.3	93	4	AAAM38501	AAAM38501 Peptide #
30	21	3.3	93	4	AAAM78238	AAAM78238 Human bon
31	20	3.1	20	5	ABB75042	ABB75042 Human lun
32	20	3.1	20	5	ABB75041	ABB75041 Human lun
33	20	3.1	20	5	ABP61961	ABP61961 Human lun
34	20	3.1	20	5	ABP61962	ABP61962 Human lun
35	20	3.1	20	7	ADA28504	ADA28504 Human lun
36	20	3.1	20	7	ADA28505	ADA28505 Human lun
37	19	3.0	171	4	AAU16166	AAU16166 Human nov
38	19	3.0	171	4	AAU16583	AAU16583 Human nov
39	19	3.0	171	6	ABU55235	ABU55235 Human nov
40	19	3.0	171	6	ABU55652	ABU55652 Human nov
41	19	3.0	556	5	ABG96346	ABG96346 Human ova
42	19	3.0	594	4	ABG06795	ABG06795 Novel hum
43	19	3.0	614	4	ABG06794	ABG06794 Novel hum
44	19	3.0	620	4	AAU16163	AAU16163 Human nov
45	19	3.0	620	6	ABU55232	ABU55232 Human nov

ALIGNMENTS

RESULT 1

AAU16161

ID AAU16161 standard; protein; 261 AA.

XX AAU16161;

AC AAU16161;

XX 07-NOV-2001 (first entry)

XX Human novel secreted protein, Seq ID 1114.

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; vulnerary; secreted protein; rheumatoid arthritis; hyperproliferative disorder; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder; corneal infection; wound healing; epithelial cell proliferation; skin ageing; food additive; preservative; antiproliferative.

XX Homo sapiens.

XX WC200155322-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001341.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215113P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217486P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225457P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254037P.
PR 05-JAN-2001; 2000US-0254978P.
PR XX
PR PA (HUMA-) HUMAN GENOME SCI INC.
PR Rosen CA, Barash SC, Ruben SM;
PR PI
PR XX
PR XX
PR DR WPI; 2001-488793/53.
PR DR N-PSDB; AAS26148.
PR XX
PR PT New nucleic acid molecules encoding 461 human secreted proteins for
PR PT diagnosing, preventing, treating or ameliorating medical conditions and
PR PT used as food additives or preservatives.
PR XX
PR XX Claim 11; SEQ ID NO 1114; 980pp; English.
PR PS The invention relates to isolated nucleic acid molecules and their
PR XX encoded secreted proteins. The nucleic acids and proteins are used to
PR CC

CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Alignment Scores:
 Pred. No.: 2 63e-200 Length: 261
 Score: 216.00 Matches: 216
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.80% Indels: 0
 DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x AAU16161 (1-261)

QY	951	TCTCAGTATCCCTGGCTGAACCTGGCTGCTGAGGTCTTTTCCAGCTTCATCCAGC	1010
Db	46	SerHisLeuileProGlyLeuAsnLeuAlaValGlyLeuPheProAlaSerSer	65
QY	1011	GCAGTCCCGCGCTCCACAGCGTACTGCGGCTGCTCCCTATAGCTCTTTATGACG	1070
Db	66	AlaValProProProSerSerValThrGlyAlaAlaProTyrSerPheMetGln	85
QY	1071	GCTCCGAGCAGAGATGTCAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATC	1130
Db	86	AlaProGluGlnGluMetValGlnValPheileProAlaGlnAlaValGlyAlaile	105
QY	1131	GGCAGAAGGGGAGCAGATCAACAGCTCTCCGCTTGGCCAGCGCTCCATCAAGAT	1190
Db	106	GlyLysLysGlyGlnHisileLysGlnLeuSerArgPheAlaSerileLysile	125
QY	1191	GCACCCCGAARACCTGACTCCAAAGTTCTGATGTTATCATCTGACCGCCAGAG	1250
Db	126	AlaProGluThrProAspSerLysValArgMetValileileThrGlyProGlu	145
QY	1251	GCCCAATTCAGGCTCGGGAAGAACTATGGCAAACTCAAGAGGAGAACTCTTTGAT	1310
Db	146	AlaGlnPheLysAlaGlnGlyArgileTyrGlyLysLeuLysGluAsnPhePheGly	165
QY	1311	CCCAAGGAGAGTGAAGTGGAGACCCACATACGTGTGCGCAGCATCAGCGTGGCGG	1370
Db	166	ProLysGluGlnValLysLeuGluThrHisileArgValProAlaSerAlaAlaGlyArg	185
QY	1371	GTCATTGGCAAGGTGGAAGAAACGGTGAACAGTTCAGATTGACCGCAGCTGAGGTG	1430
Db	186	ValileGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal	205
QY	1431	GTAGTACCAAGAGACGACCCCTGATGAGAACGACAGCTCATCTGGAATCATCGGA	1490
Db	206	ValValProArgaspGlnThrProAspGluAsnAspGlnValileLysileGly	225
QY	1491	CATTCTATCCAGTTCAGATGGTCAACGGAAGATCCGACATCTCGCCAGGTTAAG	1550
Db	226	HisPheTyrAlaSerGlnMetAlaGlnArgLysileArgaspileLeuAlaGlnValLys	245

QY	1551	CAGCAGCATCAGAGGACAGAGTAAACAGGCCAGCCAGCAGGAGGAG 1598
Db	246	GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 261
RESULT 2		
ABU5230		
ID	ABU5230	standard; protein; 261 AA.
XX	AC	ABU5230;
XX	DT	18-MAR-2003 (first entry)
XX	DE	Human novel polypeptide #317.
XX	KW	Human; neural disorder; immune system disorder; renal disorder;
XX	KW	muscular disorder; respiratory disease; reproductive disorder;
XX	KW	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX	KW	hyperproliferative disorder; inflammatory disease; allergic reaction;
XX	KW	blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX	KW	cardiovascular; nephropatic; cytosolic; antiallergic; thrombolytic;
XX	KW	haemostatic; antiarteriosclerotic.
OS	XX	Homo sapiens.
XX	XX	US2002132753-A1.
XX	XX	19-SEP-2002.
XX	PF	17-JAN-2001; 2001US-00764864.
XX	XX	31-JAN-2000; 2000US-0179065P.
PR	PR	04-FEB-2000; 2000US-0180628P.
PR	PR	28-JUN-2000; 2000US-0214886P.
PR	PR	07-JUL-2000; 2000US-0216647P.
PR	PR	07-JUL-2000; 2000US-0216880P.
PR	PR	11-JUL-2000; 2000US-0217487P.
PR	PR	11-JUL-2000; 2000US-0217496P.
PR	PR	14-JUL-2000; 2000US-0218290P.
PR	PR	26-JUL-2000; 2000US-0220963P.
PR	PR	26-JUL-2000; 2000US-0220964P.
PR	PR	14-AUG-2000; 2000US-0224518P.
PR	PR	14-AUG-2000; 2000US-0224519P.
PR	PR	14-AUG-2000; 2000US-0225267P.
PR	PR	14-AUG-2000; 2000US-0225268P.
PR	PR	14-AUG-2000; 2000US-0225270P.
PR	PR	14-AUG-2000; 2000US-0225447P.
PR	PR	14-AUG-2000; 2000US-0225757P.
PR	PR	14-AUG-2000; 2000US-0225758P.
PR	PR	22-AUG-2000; 2000US-0226868P.
PR	PR	30-AUG-2000; 2000US-0228924P.
PR	PR	01-SEP-2000; 2000US-0229287P.
PR	PR	01-SEP-2000; 2000US-0229343P.
PR	PR	01-SEP-2000; 2000US-0229344P.
PR	PR	01-SEP-2000; 2000US-0229345P.
PR	PR	05-SEP-2000; 2000US-0229509P.
PR	PR	05-SEP-2000; 2000US-0229513P.
PR	PR	08-SEP-2000; 2000US-0231413P.
PR	PR	21-SEP-2000; 2000US-0234223P.
PR	PR	21-SEP-2000; 2000US-0234274P.
PR	PR	25-SEP-2000; 2000US-0234997P.
PR	PR	27-SEP-2000; 2000US-0235834P.
PR	PR	29-SEP-2000; 2000US-0236327P.
PR	PR	29-SEP-2000; 2000US-0236367P.
PR	PR	29-SEP-2000; 2000US-0236368P.
PR	PR	29-SEP-2000; 2000US-0236369P.
PR	PR	29-SEP-2000; 2000US-0236370P.
PR	PR	02-OCT-2000; 2000US-0236802P.
PR	PR	02-OCT-2000; 2000US-0237037P.
PR	PR	02-OCT-2000; 2000US-0237038P.
PR	PR	02-OCT-2000; 2000US-0237039P.
PR	PR	02-OCT-2000; 2000US-0237040P.
PR	PR	13-OCT-2000; 2000US-0239935P.
PR	PR	20-OCT-2000; 2000US-0240960P.

Db	354 AspValAlaMetSer 359	PR	06-SEP-2000;	2000US-0230438P.
RESULT 5		PR	08-SEP-2000;	2000US-0231242P.
AAU16579		PR	08-SEP-2000;	2000US-0231243P.
ID AAU16579 standard; protein; 250 AA.		PR	08-SEP-2000;	2000US-0231244P.
XX AC AAU16579;		PR	08-SEP-2000;	2000US-0231413P.
XX DT 07-NOV-2001 (first entry)		PR	08-SEP-2000;	2000US-0231414P.
XX DE Human novel secreted protein, Seq ID 1532.		PR	08-SEP-2000;	2000US-0232080P.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;		PR	08-SEP-2000;	2000US-0232081P.
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;		PR	14-SEP-2000;	2000US-0232397P.
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;		PR	14-SEP-2000;	2000US-0232398P.
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;		PR	14-SEP-2000;	2000US-0232399P.
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;		PR	14-SEP-2000;	2000US-0232400P.
KW cerebral ischaemia; angiogenesis; nervous system disorder;		PR	14-SEP-2000;	2000US-0232401P.
KW Alzheimer's disease; infection; ocular disorder; corneal infection;		PR	14-SEP-2000;	2000US-0233063P.
KW wound healing; epithelial cell proliferation; skin ageing; food additive;		PR	14-SEP-2000;	2000US-0233064P.
KW preservative; antiproliferative.		PR	14-SEP-2000;	2000US-0233065P.
XX OS Homo sapiens.		PR	21-SEP-2000;	2000US-0234223P.
XX PN WO200155322-A2.		PR	21-SEP-2000;	2000US-0234224P.
XX PD 02-AUG-2001.		PR	25-SEP-2000;	2000US-0234937P.
XX PF 17-JAN-2001; 2001WO-US001341.		PR	25-SEP-2000;	2000US-0234938P.
PR 31-JAN-2000; 2000US-0179065P.		PR	26-SEP-2000;	2000US-0235484P.
PR 04-FEB-2000; 2000US-0180628P.		PR	27-SEP-2000;	2000US-0235834P.
PR 24-FEB-2000; 2000US-0184664P.		PR	27-SEP-2000;	2000US-0235836P.
PR 02-MAR-2000; 2000US-0186350P.		PR	29-SEP-2000;	2000US-0236337P.
PR 16-MAR-2000; 2000US-0189874P.		PR	29-SEP-2000;	2000US-0236367P.
PR 17-MAR-2000; 2000US-0190076P.		PR	29-SEP-2000;	2000US-0236368P.
PR 18-APR-2000; 2000US-0198123P.		PR	29-SEP-2000;	2000US-0236370P.
PR 19-MAY-2000; 2000US-0205515P.		PR	02-OCT-2000;	2000US-0236602P.
PR 07-JUN-2000; 2000US-0209457P.		PR	02-OCT-2000;	2000US-0237037P.
PR 28-JUN-2000; 2000US-0214886P.		PR	02-OCT-2000;	2000US-0237038P.
PR 30-JUN-2000; 2000US-0215135P.		PR	02-OCT-2000;	2000US-0237039P.
PR 07-JUL-2000; 2000US-0216647P.		PR	02-OCT-2000;	2000US-0237040P.
PR 07-JUL-2000; 2000US-0216880P.		PR	13-OCT-2000;	2000US-0239935P.
PR 11-JUL-2000; 2000US-0217487P.		PR	13-OCT-2000;	2000US-0239937P.
PR 11-JUL-2000; 2000US-0217496P.		PR	20-OCT-2000;	2000US-0240960P.
PR 14-JUL-2000; 2000US-0218230P.		PR	20-OCT-2000;	2000US-0241221P.
PR 26-JUL-2000; 2000US-0220963P.		PR	20-OCT-2000;	2000US-0241785P.
PR 26-JUL-2000; 2000US-0220964P.		PR	20-OCT-2000;	2000US-0241786P.
PR 14-AUG-2000; 2000US-0224518P.		PR	20-OCT-2000;	2000US-0241808P.
PR 14-AUG-2000; 2000US-0224519P.		PR	20-OCT-2000;	2000US-0241809P.
PR 14-AUG-2000; 2000US-0225213P.		PR	01-NOV-2000;	2000US-0244617P.
PR 14-AUG-2000; 2000US-0225214P.		PR	08-NOV-2000;	2000US-0246474P.
PR 14-AUG-2000; 2000US-0225256P.		PR	08-NOV-2000;	2000US-0246475P.
PR 14-AUG-2000; 2000US-0225257P.		PR	08-NOV-2000;	2000US-0246476P.
PR 14-AUG-2000; 2000US-0225268P.		PR	08-NOV-2000;	2000US-0246477P.
PR 14-AUG-2000; 2000US-0225270P.		PR	08-NOV-2000;	2000US-0246478P.
PR 14-AUG-2000; 2000US-0225447P.		PR	08-NOV-2000;	2000US-0246523P.
PR 14-AUG-2000; 2000US-0225757P.		PR	08-NOV-2000;	2000US-0246524P.
PR 14-AUG-2000; 2000US-0225758P.		PR	08-NOV-2000;	2000US-0246525P.
PR 14-AUG-2000; 2000US-0225759P.		PR	08-NOV-2000;	2000US-0246526P.
PR 18-AUG-2000; 2000US-0226279P.		PR	08-NOV-2000;	2000US-0246527P.
PR 22-AUG-2000; 2000US-0226681P.		PR	08-NOV-2000;	2000US-0246528P.
PR 22-AUG-2000; 2000US-0226686P.		PR	08-NOV-2000;	2000US-0246532P.
PR 22-AUG-2000; 2000US-0227182P.		PR	08-NOV-2000;	2000US-0246609P.
PR 23-AUG-2000; 2000US-0227009P.		PR	08-NOV-2000;	2000US-0246610P.
PR 30-AUG-2000; 2000US-0228924P.		PR	08-NOV-2000;	2000US-0246611P.
PR 01-SEP-2000; 2000US-0229287P.		PR	17-NOV-2000;	2000US-0246613P.
PR 01-SEP-2000; 2000US-0229343P.		PR	17-NOV-2000;	2000US-0249207P.
PR 01-SEP-2000; 2000US-0229344P.		PR	17-NOV-2000;	2000US-0249208P.
PR 01-SEP-2000; 2000US-0229345P.		PR	17-NOV-2000;	2000US-0249209P.
PR 05-SEP-2000; 2000US-0229509P.		PR	17-NOV-2000;	2000US-0249210P.
PR 05-SEP-2000; 2000US-0229513P.		PR	17-NOV-2000;	2000US-0249211P.
PR 06-SEP-2000; 2000US-0230437P.		PR	17-NOV-2000;	2000US-0249212P.
		PR	17-NOV-2000;	2000US-0249213P.
		PR	17-NOV-2000;	2000US-0249214P.
		PR	17-NOV-2000;	2000US-0249215P.
		PR	17-NOV-2000;	2000US-0249216P.
		PR	17-NOV-2000;	2000US-0249217P.
		PR	17-NOV-2000;	2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 FI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-488783/53.
 DR N-PSDB; AAS26566.
 XX
 PT New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX
 PS Claim 11; SEQ ID NO 1532; 980pp; English.
 XX

CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Alignment Scores:
 Pred. No.: 3 41e-103 Length: 250
 Score: 116.00 Matches: 116
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.15% Indels: 0
 DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x AAU16579 (1-250)

QY 1251 GCCAATTCAGGCTCAGGCAAGAACTATGTCAGAACTCAAGCAGAGAACTCTTGGT 1310
 |||||
 Db 135 AlaGlnPheLysAlaGlnGlyArgGlyLeuLysGlnGluAsnPhePheGly 154

QY 1311 CCCAAGGAGGAAGTGAAGTGGAGACCCACATACGTGTGCAGCATCAGAGCTGGCCGG 1370
 |||||
 Db 155 ProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArg 174
 |||||
 QY 1371 GTCATTGGCAAGGTGGAACAAACCGTGAACGAGTTGCAGAAATTTGACCGCAGCTGAGGTG 1430
 |||||
 Db 175 ValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 194
 |||||
 QY 1431 GTAGTACCAAGACACACCCCTGATGAGAACCCAGGTCATCGTGAATAATCATCGGA 1490
 |||||
 Db 195 ValValProArgAspGlnThrProAspGluAsnAspGlnValIleValIleGly 214
 |||||
 QY 1491 CATTTCTATGCTCAGATGGCTCAACGGAGATCCGAGACATCTCGGCCAGCTTAAG 1550
 |||||
 Db 215 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAlaGlnValLys 234
 |||||
 QY 1551 CAGCAGCATCAGAGGACAGAGTAACAGGCCCCAGCAGCAGGAGAG 1598
 |||||
 Db 235 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 250
 |||||
 RESULT 6
 ASU55648
 ID ASU55648 standard; protein; 250 AA.
 XX
 AC ASU55648;
 XX
 DT 18-MAR-2003 (first entry)
 XX
 DE Human novel polypeptide #735.
 XX
 KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 PN US2002132753-A1.
 XX
 PD 19-SEP-2002.
 XX
 PF 17-JAN-2001; 2001US-00764864.
 XX
 PP 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229503P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241805P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251858P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-147444/14.
 DR N-PSDB; ABX73907.
 XX
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1532; 402pp; English.
 XX
 PS The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX
 SQ Sequence 250 AA;
 XX
 Alignment Scores:
 Pred. No.: 3.41e-103 Length: 250
 Score: 116.00 Matches: 116
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.15% Indels: 0
 DB: 6 Gaps: 0
 US-09-270-437D-7 (1-1946) x ABU55648 (1-250)
 QY 1251 GCCCAATTCAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGAGGAACTTTTGTG 1310
 Db 135 AlaGlnPhelysAlaGlnGlyArgIleTyrGlyLeuLysGluGlnPheGly 154
 QY 1311 CCCAAGGAGGAAGTGAAGCTGGAGACCCATACATAGTGTGCCACATCAGCAGCTGCCCG 1370

Db 155 ProLysGluGluValValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArg 174
 QY 1371 GTCAATTGCAAGGTGGAAAACCGTGAAACGAGTTGCAAGATTTCAGCGCAGCTGAGGTG 1430
 Db 175 ValIleGlyLysGlyGlyThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 194
 QY 1431 GTAGTACCAAGAGACACGACCCCTGATGAGAACGACACGAGGTCTCATCGTGAATAATCATCCGA 1490
 Db 195 ValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLysIleIleGly 214
 QY 1491 CATTTCATGCGAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCAGGTTAAG 1550
 Db 215 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAlaGlnValLys 234
 QY 1551 CAGCAGCATCAGAGGGACAGAGTAACACGAGCCCGACGAGGAGGAAG 1598
 Db 235 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 250
 RESULT 7
 AAB11328
 ID AAB11328 standard; protein; 579 AA.
 XX AAB11328;
 XX
 DT 21-FEB-2001 (first entry)
 XX
 DE Human lung cancer-associated protein L523S.
 XX
 KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
 KW vaccine; detection.
 OS Homo sapiens.
 XX
 FN WO2000061612-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 03-APR-2000; 2000WO-US008896.
 XX
 PR 02-APR-1999; 99US-00285479.
 PR 17-DEC-1999; 99US-00466396.
 PR 30-DEC-1999; 99US-00476496.
 PR 10-JAN-2000; 2000US-00480884.
 PR 22-FEB-2000; 2000US-00510376.
 XX
 PA (CORI-) CORIXA CORP.
 PI Wang T, Fan L;
 XX
 WIPI; 2000-628399/60.
 DR N-PSDB; AAC65900.
 XX
 PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
 PT protein is used for detecting and monitoring progression of lung cancer
 PT in a patient.
 XX
 PS Claim 3; Page 186-188; 261pp; English.
 XX
 CC This invention describes a novel isolated polypeptide (I) which
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
 CC which have cytostatic activity. The polypeptides and polynucleotides are
 CC used in compositions and vaccines to inhibit the development of cancer,
 CC especially lung cancer, in a patient. Methods described in the invention
 CC can be used to monitor the progression of a cancer by carrying out the
 CC detection at subsequent time points and comparing the results from the
 CC different time points. CD4+ and/or CD8+ T-cells isolated from a patient
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting
 CC cells expressing P2 and then administered to the patient to inhibit
 CC development of cancer
 XX
 SQ Sequence 579 AA;

Alignment Scores:

Pred. No.:	1.43e-25	Length:	579
Score:	36.00	Matches:	36
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	3	Gaps:	0

US-09-270-437D-7 (1-1946) x AAB11328 (1-579)

QY	492	GTGGTGCCATTATTCGCAAGAGGGGCGCACCATCCGCAACATCAAAACAGACCCAG	551
Db	207	ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln	226
QY	552	TCCAAGATAGCTGCATAGGAGGAGAACGCGAGTGCAGCTCAAAA	599
Db	227	SerLysIleAspValHisArgLysGluAsnAlaGlyAlaGluLys	242

RESULT 8

ID	AAB11365	standard; protein; 579 AA.
XX	AC	AAB11365;
XX	DT	21-FEB-2001 (first entry)
XX	DE	Human lung cancer associated antigen L523S.
XX	DE	Human lung cancer; treatment; human; tumor; immunogenic; cytostatic;
XX	KW	lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX	KW	vaccine; detection.
XX	OS	Homo sapiens.
XX	XX	WO200061612-A2.
XX	PN	19-OCT-2000.
XX	PD	03-APR-2000; 2000WO-US008896.
XX	PF	02-APR-1999; 99US-00285479.
XX	PR	17-DEC-1999; 99US-00466396.
XX	PR	30-DEC-1999; 99US-00476496.
XX	PR	10-JAN-2000; 2000US-00480884.
XX	PR	22-FEB-2000; 2000US-00510376.
XX	PA	(CORI-) CORIXA CORP.
XX	PI	Wang T, Fan L;
XX	PI	WPI; 2000-628399/60.
XX	DR	N-PSDB; AAC66035.
XX	PT	Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX	PT	protein is used for detecting and monitoring progression of lung cancer
XX	PT	in a patient.
XX	PS	Claim 3; Page 259-261; 261pp; English.
XX	CC	This invention describes a novel isolated polypeptide (I) which
XX	CC	comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX	CC	which have cytostatic activity. The polypeptides and polynucleotides are
XX	CC	used in compositions and vaccines to inhibit the development of cancer,
XX	CC	especially lung cancer, in a patient. Methods described in the invention
XX	CC	can be used to monitor the progression of a cancer by carrying out the
XX	CC	detection at subsequent time points and comparing the results from the
XX	CC	different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX	CC	are treated with P2, polynucleotides encoding P2 or antigen presenting
XX	CC	cells expressing P2 and then administered to the patient to inhibit
XX	CC	development of cancer
XX	SQ	Sequence 579 AA;

Alignment Scores:

Pred. No.: 1.43e-25 Length: 579

Score: 36.00 Matches: 36

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.63% Indels: 0

DB: 3 Gaps: 0

US-09-270-437D-7 (1-1946) x AAB11365 (1-579)

QY	492	GTGGTGCCATTATTCGCAAGAGGGGCGCACCATCCGCAACATCAAAACAGACCCAG	551
Db	207	ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln	226
QY	552	TCCAAGATAGCTGCATAGGAGGAGAACGCGAGTGCAGCTCAAAA	599
Db	227	SerLysIleAspValHisArgLysGluAsnAlaGlyAlaGluLys	242

RESULT 9

ID	ABB74960	standard; protein; 579 AA.
XX	AC	ABB74960;
XX	DT	01-MAY-2002 (first entry)
XX	DE	Human lung tumour L523S protein sequence SEQ ID NO:176.
XX	DE	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX	KW	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX	KW	immune response.
XX	OS	Homo sapiens.
XX	XX	WO200200174-A2.
XX	PN	03-JAN-2002.
XX	PD	28-JUN-2001; 2001WO-US021065.
XX	PF	28-JUN-2000; 2000US-00606421.
XX	PR	02-AUG-2000; 2000US-00630940.
XX	PR	21-AUG-2000; 2000US-00643597.
XX	PR	15-SEP-2000; 2000US-00662786.
XX	PR	09-OCT-2000; 2000US-00685696.
XX	PR	12-DEC-2000; 2000US-00735705.
XX	PR	07-MAY-2001; 2001US-00850716.
XX	PA	(CORI-) CORIXA CORP.
XX	PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX	PI	McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX	PI	Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX	XX	WPI; 2002-090513/12.
XX	DR	N-PSDB; ABL49119.
XX	PT	Polynucleotides encoding lung tumor polypeptides, useful for treating
XX	PT	lung cancer or stimulating an immune response.
XX	PS	Example 2; Page 267-268; 374pp; English.
XX	CC	The present invention describes human lung tumour proteins. Human lung
XX	CC	tumour proteins and polynucleotides have cytostatic and immunostimulant
XX	CC	activities, and can be used in vaccine production. Compositions
XX	CC	comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX	CC	proteins, T cell populations, or antigen presenting cells that express
XX	CC	the lung tumour proteins are useful for treating lung cancer or
XX	CC	stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX	CC	ABB75070 represent sequences used in the exemplification of the present
XX	CC	invention
XX	SQ	Sequence 579 AA;

Alignment Scores:

```
Pred. No.: 1.43e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 5 Gaps: 0

US-09-270-437D-7 (1-1946) x ABB74960 (1-579)
QY 492 GTGGTGCCTATTTCGCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAG 551
DB 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 552 TCCAAGATAGACGTGCATAGGAGGAGAGCGAGGTGCAGCTGAAAAA 599
DB 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 10
ABB75053
ID ABB75053 standard; protein; 579 AA.
AC ABB75053;
XX
XX 01-MAY-2002 (first entry)
XX
XX Human lung tumour L523S recombinant protein sequence SEQ ID NO:446.
XX
XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX immune response.
XX
XX Homo sapiens.
XX
XX WO200200174-A2.
XX
XX 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US021065.
XX
XX 28-JUN-2000; 2000US-00606421.
XX
XX 02-AUG-2000; 2000US-00630940.
XX
XX 21-AUG-2000; 2000US-00643597.
XX
XX 15-SEP-2000; 2000US-00662786.
XX
XX 09-OCT-2000; 2000US-00685696.
XX
XX 12-DEC-2000; 2000US-00735705.
XX
XX 07-MAY-2001; 2001US-00850716.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
XX WPI; 2002-090513/12.
XX
XX N-PSDB; ABL49297.
XX
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response.
XX
XX Claim 2; Page 365-367; 374pp; English.
XX
XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX proteins, T cell populations, or antigen presenting cells that express
XX the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention
XX
XX Sequence 579 AA;
XX
XX Alignment Scores:

Pred. No.: 1.43e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 5 Gaps: 0

US-09-270-437D-7 (1-1946) x ABB75053 (1-579)
QY 492 GTGGTGCCTATTTCGCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAG 551
DB 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 552 TCCAAGATAGACGTGCATAGGAGGAGAGCGAGGTGCAGCTGAAAAA 599
DB 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 11
ABB74997
ID ABB74997 standard; protein; 579 AA.
AC ABB74997;
XX
XX 01-MAY-2002 (first entry)
XX
XX Human lung tumour L523S protein sequence SEQ ID NO:348.
XX
XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX immune response.
XX
XX Homo sapiens.
XX
XX WO200200174-A2.
XX
XX 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US021065.
XX
XX 28-JUN-2000; 2000US-00606421.
XX
XX 02-AUG-2000; 2000US-00630940.
XX
XX 21-AUG-2000; 2000US-00643597.
XX
XX 15-SEP-2000; 2000US-00662786.
XX
XX 09-OCT-2000; 2000US-00685696.
XX
XX 12-DEC-2000; 2000US-00735705.
XX
XX 07-MAY-2001; 2001US-00850716.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
XX WPI; 2002-090513/12.
XX
XX N-PSDB; ABL49254.
XX
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response.
XX
XX Example 2; Page 330-332; 374pp; English.
XX
XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX proteins, T cell populations, or antigen presenting cells that express
XX the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention
XX
XX Sequence 579 AA;
XX
XX Alignment Scores:
```


SQ Sequence 579 AA;
 CC invention
 XX
 SQ Sequence 579 AA;
 Alignment Scores:
 Pred. No.: 1.43e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 5 Gaps: 0
 US-09-270-437D-7 (1-1946) x ABP61917 (1-579)
 QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCAAAACAGACCCAG 551
 Db 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAGATAGACGTGCATAGAGAGGAGACGCGAGTGCAGCTCAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242
 RESULT 14
 ABP61974
 ID ABP61974 standard; protein; 579 AA.
 XX
 AC ABP61974;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human lung cancer associated protein sequence SEQ ID NO:449.
 XX
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200247534-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 30-NOV-2001; 2001WO-US047576.
 XX
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
 DR WPI: 2002-583465/62.
 DR N-PSDB; ABQ92485.
 XX
 Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 the polynucleotides, useful in pharmaceutical compositions such as
 vaccines and as markers to indicate the presence of lung cancer.
 PT
 PS Claim 9; Page 375-377; 381pp; English.
 XX
 The present invention describes isolated human lung carcinoma
 polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 activity, and can be used in gene therapy and in vaccines. Compositions
 comprising (I) or (II) can be used for stimulating an immune response in
 a patient and for treating lung cancer in a patient. Oligonucleotides of
 (I) can be used for detecting the presence of a cancer in a patient, by
 obtaining a biological sample from the patient, contacting the biological
 sample with the oligonucleotide, detecting in the sample, an amount of
 polynucleotide that hybridises to the oligonucleotide and comparing the
 amount of polynucleotide that hybridises to the oligonucleotide to a
 predetermined cut-off value, and determining the presence of a cancer in
 the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 vaccines. (I) is useful as a marker to indicate the presence or absence
 of a cancer such as lung cancer. ABQ92145 to ABQ92485 and ABP61866 to

CC ABP61992 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 579 AA;
 Alignment Scores:
 Pred. No.: 1.43e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 5 Gaps: 0
 US-09-270-437D-7 (1-1946) x ABP61974 (1-579)
 QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCAAAACAGACCCAG 551
 Db 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAGATAGACGTGCATAGAGAGGAGACGCGAGTGCAGCTCAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242
 RESULT 15
 ABP61880
 ID ABP61880 standard; protein; 579 AA.
 XX
 AC ABP61880;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human lung cancer associated protein sequence SEQ ID NO:176.
 XX
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200247534-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 30-NOV-2001; 2001WO-US047576.
 XX
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
 DR WPI: 2002-583465/62.
 DR N-PSDB; ABQ92305.
 XX
 Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 the polynucleotides, useful in pharmaceutical compositions such as
 vaccines and as markers to indicate the presence of lung cancer.
 PT
 PS Example 2; Page 274-275; 381pp; English.
 XX
 The present invention describes isolated human lung carcinoma
 polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 activity, and can be used in gene therapy and in vaccines. Compositions
 comprising (I) or (II) can be used for stimulating an immune response in
 a patient and for treating lung cancer in a patient. Oligonucleotides of
 (I) can be used for detecting the presence of a cancer in a patient, by
 obtaining a biological sample from the patient, contacting the biological
 sample with the oligonucleotide, detecting in the sample, an amount of
 polynucleotide that hybridises to the oligonucleotide and comparing the
 amount of polynucleotide that hybridises to the oligonucleotide to a
 predetermined cut-off value, and determining the presence of a cancer in
 the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 vaccines. (I) is useful as a marker to indicate the presence or absence
 of a cancer such as lung cancer. ABQ92145 to ABQ92485 and ABP61866 to

CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (I) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. ABQ2145 to ABQ2486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention

XX SQ Sequence 579 AA;

Alignment Scores:
 Pred. No.: 1.43e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 5 Gaps: 0

US-09-270-437D-7 (1-1946) x ABP61880 (1-579)

OY	492	GTGGGTGCATTATTGGCAAGGAGGGCCACCATCCGACATCACAACACAGACCCAG	551
DB	207	ValGlyAlaIleileGlyGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln	226
OY	552	TCCAAGATAGACGTGCATAGGAGGAGAACGCGAGTGCAGCTGAAAAA	599
DB	227	SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys	242

Search completed: August 6, 2004, 13:20:29
 Job time : 95.4096 secs

Result No.	Score	Query *		Length	DB	ID	Description
		Match	Match				
1	166	26	0	577	3	US-09-361-855-2	Sequence 2, Appl
2	49	7.7	49	3	US-09-361-855-18	US-09-361-855-20	Sequence 18, Appl
3	48	7.5	48	3	US-09-361-855-20	US-09-361-855-17	Sequence 20, Appl
4	47	7.4	47	3	US-09-361-855-17	US-09-361-855-19	Sequence 17, Appl
5	45	7.0	47	3	US-09-361-855-19	US-09-643-597-176	Sequence 19, Appl
6	36	5.6	579	4	US-09-643-597-176	US-09-643-597-348	Sequence 176, Appl
7	36	5.6	579	4	US-09-643-597-348	US-09-480-8848-176	Sequence 348, Appl
8	36	5.6	579	4	US-09-480-8848-176	US-09-480-8848-176	Sequence 176, Appl
9	36	5.6	579	4	US-09-480-8848-176	US-09-480-8848-176	Sequence 176, Appl
10	36	5.6	579	4	US-09-480-8848-176	US-09-480-8848-176	Sequence 176, Appl
11	36	5.6	579	4	US-09-480-8848-176	US-09-480-8848-176	Sequence 176, Appl
12	36	5.6	579	4	US-09-480-8848-176	US-09-480-8848-176	Sequence 176, Appl

דב
274 CINGULAR

דב
274 CINGULAR

Db 214 GluGlyAlaThrIleA

274 57

453	GTGGACATCCCCCTTCGGCTCCTGTCGCCACACCAAGTATGGTGCCATTATTTGCCAAG	512
Qy		
194	ValAspIleProLeuArgLeuLeuValProThrGlnrValGlyAlaIleIleGlyLys	213
Db		
513	GAGGGGGGCCCATTCGCGAATCACAAAACAGACCCAGTCCAGATAGACGTGCATATAGG	572
Qy		
214	GluGlyValThrIleArgAsnIleThrIysGlnThrGlnSerLysIleAspValHisArg	233
Db		

QY	573	AAGGAGAAACG	CAGGTGCACGCTG	AAAAACCCATCAGTGTGCAC	TCCACCCCTGAGGGCTGC	632
Db	234	LysGluAsnAla	glaGluLysAla	ileSerValHisSer	ThrProGluGlyCys	253
QY	633	TCTCCGCTTGA	TAAAGATGATCTTG	GAGATTATGCATAAAGAGG	CGCTAAGGACACCAAAACG	692
Db	254	SerSerAlaCys	lysSweileuGlu	ileMetHisLysGlu	alLysAspThrLysThr	273
QY	693	GCTGACGAGGT	TCCTCTGAAGATCCTG	CCCATCTGCCCATATAACTTT	GTAGGCGCTCATGGC	752
Db	274	AlaAspGluVal	ProLeuLysileLeu	AlaHisAsnAsnPheVal	GlyAsGLeuIleGly	293
QY	753	AGGAGGAGCGA	ACTGARGAGGTGAGCA	AGTACAGCAATACCCGAG	ACAAAATCACCANTCC	812
Db	294	LysGluGlyArg	AsnLeuLysLysVal	GluGlnAspThrGlu	thrLysIleThrIleSer	313
QY	813	TCGTTGCAAGAC	CTTACCCTTTTACAA	CCCTGAGAGGACCATCTGT	GTAAGGGGGCCATC	872
Db	314	SerLeuGlnAsp	LeuThrLeuTyra	snProGluArgThrIle	thrValLysGlyAlaIle	333
QY	873	GAGATTTGTCAG	GCGCCGACGCAATAT	ATGAGAAAGTTCGGAGG	CGCTATGAGAAT	932
Db	334	GluAsnCysCys	ArgAlaGluGlnGlu	ileMetLysLysVal	ArgGluAlaTyrGlu	Asn 353
QY	933	GATGTGGCTGCC	ATGACG	950		
Db	354	AspValAlaAla	MetSer	359		

```

RESULT 2
US-09-261-855-18
: Sequence 18, Application US/09261855A
: Patent No. 6255055
: GENERAL INFORMATION:
: APPLICANT: Ross, Jeffrey
: TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
: TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

```

US-09-270-437D-7 (1-1946) x US-09-261-855-18 (1-49)

QY	714	ATCTGGGCCATAATAACTTTGTAGGCGCTCTATTGGCAAGGAAGACGACCTGAAG	7773
Db	1	l l l l l e u a l a h s a s n s p h e v a l g l y a r g l e u l e g l y l y s g l u g l y a r g a n l e u l y s	20
QY	774	AAGGTAGACGAAGATACCGAGACACAAAATCACCATCTCTCTGTGTGCAAGACCTTACCTTT	8333
Db	21	LysValGluGlnAspThrGluThrLysLeuIleThrIleSerSerLeuGlnAspLeuThrLeu	40
QY	834	TACACCTCTGAGAGCACTACTGTG	860
Db	41	TyrAsnProGluArgThrIleThrVal	49

RESULT 3
US-09-261-855-20
; Sequence 20. Application US/09261855A

```

; Patent No. 5255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-20

Alignment Scores:
Pred. No.: 5,87e-37 Length: 48
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.51% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-261-855-20 (1-48)

Qy 1341 ATAGGTGTGCCAGCATCAGCAGCTGCGCGGTGCTATTGGCAAAAGGTGGAAAACCGTGAAC 1400
Db 1 IleArgValProAlaSerAlaAlaGlyArgValIleGlyysGlyGlyIysThrValAsn 20

Qy 1401 GAGTTGCAGAAATTGACGGCAGCTGAGGTGGTGTAGTACCAGAGACACAGACCCCTGATGAG 1460
Db 21 GluLeuGlnAsnLeuThrAlaIleAlaGluValValProArgaspGlnThrProaspGlu 40

Qy 1461 AACGACCAAGTCATCGTGAATC 1484
Db 41 AsnAspGlnValIleValIysIle 48

```

```

RESULT 4
US-09-261-855-17
; Sequence 17, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A

```


; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.: 1.12e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-480-884A-176 (1-579)

QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCGCAACATCACAAAACAGACCCAG 551
|||||
DB 207 ValGlyAlaIlelleGlyysGluGlyAlaThrIleArgAsnIleThrlyysGlnThrGln 226

QY 552 TCCAAGATAGCTGTCATAGGAGGAGAACGCGAGGTGCAGCTGAAAAA 599
|||||
DB 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 9

US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

Alignment Scores:
Pred. No.: 1.12e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-542-615A-176 (1-579)

QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCGCAACATCACAAAACAGACCCAG 551
|||||
DB 207 ValGlyAlaIlelleGlyysGluGlyAlaThrIleArgAsnIleThrlyysGlnThrGln 226

QY 552 TCCAAGATAGCTGTCATAGGAGGAGAACGCGAGGTGCAGCTGAAAAA 599
|||||
DB 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 10

US-09-542-615A-348
; Sequence 348, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-348

Alignment Scores:
Pred. No.: 1.12e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-542-615A-348 (1-579)

QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCGCAACATCACAAAACAGACCCAG 551
|||||
DB 207 ValGlyAlaIlelleGlyysGluGlyAlaThrIleArgAsnIleThrlyysGlnThrGln 226

QY 552 TCCAAGATAGCTGTCATAGGAGGAGAACGCGAGGTGCAGCTGAAAAA 599
|||||
DB 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 11

US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176

```
Alignment Scores:
Pred. No.: 1.12e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-606-421B-176 (1-579)
QY 492 GTGGGTGCCATTATTCGCAAGAGGGCGCCACCATCCGCAACATCACAAAAACAGACCCAG 551
Db 207 ValGlyAlaIleiledgLyysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 552 TCCAAATAGACGTGCATAGGAGGAGAACGCGAGTGCAGCTGAAAAA 599
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 12
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kaloe, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 1.12e-25 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-606-421B-348 (1-579)
QY 492 GTGGGTGCCATTATTCGCAAGAGGGCGCCACCATCCGCAACATCACAAAAACAGACCCAG 551
Db 207 ValGlyAlaIleiledgLyysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 552 TCCAAATAGACGTGCATAGGAGGAGAACGCGAGTGCAGCTGAAAAA 599
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 13
US-09-261-855-21
; Sequence 21, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
```

```
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-21

Alignment Scores:
Pred. No.: 6.36e-18 Length: 47
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.38% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-261-855-21 (1-47)
QY 516 GGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAG 575
Db 16 GlyAlaThrIleArgAsnIleThrLysGlnThrLysGlnThrLysGlnThrLysGlnThrLys 35
QY 576 GAGAACGCGAGTGCAGCTGAAAAA 599
Db 36 GluAsnAlaGlyAlaAlaGluLys 43

RESULT 14
US-09-261-855-24
; Sequence 24, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-24

Alignment Scores:
Pred. No.: 2.33e-09 Length: 48
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-261-855-24 (1-48)
QY 1359 CGAGTGGCGGGTCAATTCGCAAGAGTGGAAAAACGGTGAACGAGTTGCAGAAATTG 1415
Db 7 AlaAlaGlyArgValIleGlyGlyLysThrValAsnGluLeuGlnAsnLeu 25

RESULT 15
US-09-261-855-3
; Sequence 3, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-3

Alignment Scores:
Pred. No.:      0.000156      Length:      14
Score:          14.00         Matches:     14
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:      2.19%      Indels:      0
DB:                3         Gaps:        0

US-09-270-437D-7 (1-1946) x US-09-261-855-3 (1-14)

Qy      372  CGCGAGGGGCTTGGCTCTCGGGTCAGCCGCCAGGGC 413
Db      1   ArgArgGlyGlyPheGlySerArgGlyGlnProArgGlnGly 14

Search completed: August 6, 2004, 13:35:15
Job time : 24.4913 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:31:07 ; Search time 93.2249 Seconds
(without alignments)
13095.793 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 639
Sequence: 1 gctgtagcggggggctggg.....attccttcagggttttaaaa 1946

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1291235 seqs, 313682936 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2479628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlpl
-MODEL=frame+n2p.model
-O=/cgn2_1/USPTO_spool_p/US09270437/runat_06082004_141223_1936/app_query.fasta.1.5582
-DB=Published Applications AA -QPMT=fastan -SUFFIX=oligo806.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT8=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMV=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09270437@cgn_1_1_35/runat_06082004_141223_1936
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

```
1: /cgn2_6/prodata/2/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	317	49.6	577	15	US-10-313-986-500	Sequence 500, App
2	216	33.8	261	9	US-09-764-864-1114	Sequence 1114, App
3	166	26.0	577	9	US-09-873-637-2	Sequence 2, Appl1
4	116	18.2	250	9	US-09-764-864-1532	Sequence 1532, Ap
5	49	7.7	48	9	US-09-873-637-18	Sequence 18, Appl
6	48	7.5	48	9	US-09-873-637-20	Sequence 20, Appl
7	47	7.4	47	9	US-09-873-637-17	Sequence 17, Appl
8	45	7.0	47	9	US-09-873-637-19	Sequence 19, Appl
9	36	5.6	70	15	US-10-117-982-472	Sequence 472, App
10	36	5.6	70	15	US-10-313-986-472	Sequence 472, App
11	36	5.6	422	16	US-10-408-765A-2088	Sequence 2088, Ap
12	36	5.6	579	9	US-09-735-705-176	Sequence 176, App
13	36	5.6	579	9	US-09-735-705-348	Sequence 348, App
14	36	5.6	579	9	US-09-850-716A-176	Sequence 176, App
15	36	5.6	579	9	US-09-850-716A-348	Sequence 348, App
16	36	5.6	579	9	US-09-897-778-176	Sequence 176, App
17	36	5.6	579	9	US-09-897-778-348	Sequence 348, App
18	36	5.6	579	9	US-09-897-778-446	Sequence 446, App
19	36	5.6	579	9	US-09-897-778-449	Sequence 449, App
20	36	5.6	579	10	US-09-466-396A-176	Sequence 176, App
21	36	5.6	579	12	US-10-007-700-176	Sequence 176, App
22	36	5.6	579	12	US-10-007-700-348	Sequence 348, App
23	36	5.6	579	12	US-10-007-700-446	Sequence 446, App
24	36	5.6	579	12	US-10-007-700-449	Sequence 449, App
25	36	5.6	579	14	US-10-117-982-176	Sequence 176, App
26	36	5.6	579	14	US-10-117-982-348	Sequence 348, App
27	36	5.8	579	14	US-10-117-982-446	Sequence 446, App
28	36	5.6	579	14	US-10-117-982-449	Sequence 449, App
29	36	5.6	579	14	US-10-117-982-480	Sequence 480, App
30	36	5.6	579	14	US-10-117-982-484	Sequence 484, App
31	36	5.6	579	15	US-10-313-986-176	Sequence 176, App
32	36	5.6	579	15	US-10-313-986-348	Sequence 348, App
33	36	5.6	579	15	US-10-313-986-446	Sequence 446, App
34	36	5.8	579	15	US-10-313-986-449	Sequence 449, App
35	36	5.6	579	15	US-10-313-986-480	Sequence 480, App
36	36	5.6	579	15	US-10-313-986-484	Sequence 484, App
37	36	5.6	586	9	US-09-850-716A-427	Sequence 427, App
38	36	5.6	586	9	US-09-897-778-427	Sequence 427, App
39	36	5.6	586	12	US-10-007-700-427	Sequence 427, App
40	36	5.6	586	14	US-10-117-982-427	Sequence 427, App
41	36	5.6	586	15	US-10-313-986-427	Sequence 427, App
42	36	5.6	589	15	US-10-313-986-486	Sequence 486, App
43	28	4.4	47	9	US-09-873-637-21	Sequence 21, Appl
44	24	3.8	70	14	US-10-117-982-473	Sequence 473, App
45	24	3.8	70	15	US-10-313-986-473	Sequence 473, App

ALIGNMENTS

RESULT 1
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 21021.455C19
; CURRENT APPLICATION NUMBER: US/10/313.986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-500

Alignment Scores:

Pred. No.: 5,49e-293 Length: 577
 Score: 317.00 Matches: 497
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 99.60% Mismatches: 0
 Query Match: 49.61% Indels: 2
 DB: 15 Gaps: 1

US-09-270-437D-7 (1-1946) x US-10-313-986-500 (1-577)

QY 108 AGGAGCCGGAATTCCTCAATCCGAATATTCACCCAGCTCCGATCGGAAGTACTGGAC 167
 DB 79 ArgSerArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAsp 98
 QY 168 AGCTGCTGCTCAGTATGCTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGTGAAG 227
 DB 99 SerLeuLeuAlaGlnTrpGlyThrValGluAsnGlyGlnValAsnThrGluSerGlu 118
 QY 228 ACGGCACTGCTGATGCTCACTTATCCAAACCGGAGCAGACCAAGCAAGCCATCATGAAG 287
 DB 119 ThrAlaValValAsnValThrTrpSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
 QY 288 CTGAATGCCACCAAGTTGGAGAACCACTGCTGAGGCTCTTACATATCCCGATGAGCAG 347
 DB 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTrpIleProAspGluGln 158
 QY 348 ATAGCAGGAGCTGAGATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGCCCGC 407
 DB 159 IleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
 QY 408 CAGGGCTCACCTGTGGCAGCGGGGCCCAAGCAAGCAGCAGCAAGTGCACATCCCGCTT 467
 DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
 QY 468 CGGCTCTGTGTCGCCACCCAGTATGGGTGCCATTATTGGCAGAGGGGCCACCATC 527
 DB 199 ArgLeuLeuValProThrGlnTrpValGlyAlaIleIleGlyGlyGluGlyAlaThrIle 218
 QY 528 CGCAACATCAAAACAGACCCAGTCCAGATAGAGTGCATAGGAGGAGAGCAAGCAGGT 587
 DB 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
 QY 588 GAGCTGMAAAGCCATCAGTGTGCTACTCCACCCCTGAGGCTGCTCTCCGCTGTGTAAG 647
 DB 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
 QY 648 ATGATCTTGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGCTCAGCAGGTTCCT 707
 DB 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 QY 708 CTGAAGATCTGCGCCATATAACTTTGTAGGCGTCTCATTGGCAAGGAAGAGCGGAAC 767
 DB 279 LeuLysIleLeuAlaHisAsnAsnProValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 QY 768 CTGAAGAAGGTAGACAGATACCGAGACAAATAATCACCATTCTCTCGTTGCAAGACCTT 827
 DB 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
 QY 828 ACCCTTTACAACCTCAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTTGTCAGG 887
 DB 319 ThrLeuTrpAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnGlyCysArg 338
 QY 888 GCGGACGGAATATATGAGAAGTTCGGGAGGCTATGAGATGATGCTGCTCCCATG 947
 DB 339 AlaGluGlnIleMetLysLysValArgGluAlaTrpGluAsnAspValAlaAlaMet 358
 QY 948 AGC-----TCTCACCTGATCCCTGGCTGAACTGGCTGCTGAGTCTTTTCCAGCT 1001
 DB 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
 QY 1002 TCATCAGGCGAGTCCCGCGGCTCCACAGAGGTTACTGGGGCTGCTCCCTATAGTCTCC 1061
 DB 379 SerSerSerAlaValProProProProSerSerValThrGlyAlaAlaProTrpSerSer 398

QY 1062 TTTATGCAAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCCGCAGGAGTGGC 1121
 DB 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
 QY 1122 GCCATCATCGCAAGAGGGGCGACACATCAACAGCTCTCCCGTTTCCAGGCGCTCC 1181
 DB 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
 QY 1182 ATCAAGATTGCACCAACCGAAACCTGACTCCAAGTTCGATGTTATGTTATCATCTGGA 1241
 DB 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleIleThrGly 458
 QY 1242 CCGCCAGAGGCCCAATTCAGGCTCAGGAGGAATCTATGCAAACTCAAGGAGGAGAAC 1301
 DB 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTrpGlyLysLeuLysGluGluAsn 478
 QY 1302 TTCTTTGCTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
 DB 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
 QY 1362 GCTGCGCGGCTCATGCGAAAGGTGGAAACCGTGAACGAGTTGCAAGATTTGACGGCA 1421
 DB 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnLeuThrAla 518
 QY 1422 GCTCAGGTGGTGTACTCAAGAGAGACAGACCCCTGATGAGAACCAACCAAGGTTCATCGTGAAA 1481
 DB 519 AlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
 QY 1482 ATCATCGGACATTTCTATGTCAGTCAGTCAGTGCCTCAACGGAAGATCCGAGACATCTGGCC 1541
 DB 539 IleIleGlyHisPheTrpAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
 QY 1542 CAGGTTAAGCAGCAGCATCAGAGGGACAGTAACCAAGCCAGGACCGAGGACCGAGGAAG 1598
 DB 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgA-gLys 577

RESULT 2

US-09-764-864-1114
 ; Sequence 1114, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PTZ23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1114
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-864-1114

Alignment Scores:
 Pred. No.: 1,38e-196 Length: 261
 Score: 216.00 Matches: 216
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.80% Indels: 0
 DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-764-864-1114 (1-261)

QY 951 TCTCACCTGATCCCTGGCTGAACTGGCTGCTGAGTCTTTTCCAGCTTCATCCAGC 1010
 DB 46 SerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAlaSerSer 65
 QY 1011 GCAGTCCGGCGGCTCCACAGCAGGTTACTGGGGCTGCTCCCTATAGTCTCTTATGCG 1070
 DB 66 AlaValProProProProSerSerValThrGlyAlaAlaProTrpSerSerPheMetGln 85

QY 1071 GCTCCGACGACGATGCTGAGGTGTTTATCCGCCGCCAGGAGTGGGGCCCATATC 1130
DB 86 AlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValAlaIle 105
QY 1131 GGCAGAGAGGCGACGACATCAAGAGCTCTCCCGGTTTGCAGGCGCTCATCAAGATT 1190
DB 106 GlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSerIleLysIle 125
QY 1191 GCACACCCGAAACCTGACTCCAAAGTTCGTATGTTTATCATCATCTGGACCGCCAGAG 1250
DB 126 AlaProProGluThrProAspSerLysValArgMetValIleIleThrGlyProProGlu 145
QY 1251 GCCCAATTCAAGCTCAGGAGAAATCTATGCAAACTCAAGAGGAGAACTCTTTGGT 1310
DB 146 AlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPhePheGly 165
QY 1311 CCCAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1370
DB 166 ProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArg 185
QY 1371 GTCATTGGCAAGGTGCAAAACGGTGAACGAGTTGCAAGATTTCAGCGGAGCTGAGTG 1430
DB 186 ValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 205
QY 1431 GTAGTACCAAGAGACAGACCCCTGATGAGAACGACGAGTCTATGTAATCATCGGA 1490
DB 206 ValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLysIleGly 225
QY 1491 CATTTCTATGCCAGTGCATGCTCAACGGAAGTCCGAGACATCTGCGCCAGGTTAAG 1550
DB 226 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAlaGlnValLys 245
QY 1551 CAGCAGCATCAGAAGGACAGAGTACACGAGGCGCGGACCGGACCGAGGAAG 1598
DB 246 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 261

RESULT 3
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-2

Alignment Scores:
Pred. No.: 6,33e-149 Length: 577
Score: 166.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.98% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-873-637-2 (1-577)

QY 453 GTGACATCCCTTGGCTTCCTGGTCCGCCACCATGATGTGGGTCCCATTTATGGCAAG 512
DB 194 ValAspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCACCATCGCAATCACAACACAGCCAGTCCCAAGATAGACGTGCATAGG 572
DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233

QY 573 AAGGAGAACCCAGGTGAGCTGAAAAAGCCATCAGTGTGACTCCACCCCTGAGGCGCTGC 632
DB 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCys 253
QY 633 TCCTCCGCTTGAAGATGATCTTGGAGATTATCATAAAGAGGCTAAGACACACAAAACG 692
DB 254 SerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThr 273
QY 693 GCTGACAGAGTTCCTCCGAAAGATCCCTGGCCCATTAATAAATTTGTAGGCGCTCTCATTTGGC 752
DB 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGAGCGGACCTGGAAGAGGTAGACCAAGATACCGACACAAAATACCATCTCC 812
DB 294 LysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIleSer 313
QY 813 TCGTGTCAAGACCTTACCTTTTACACCTGAGAGGACCATCTGTGAAGGGGGCCATC 872
DB 314 SerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIle 333
QY 873 GAGAAATTGTCAGGCGGAGCAGGAGGAATATGAAGAAAGTTCCGGAGGCGCTATGAGAAT 932
DB 334 GluAsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsn 353
QY 933 GATGTGCTGCCATGAGC 950
DB 354 AspValAlaAlaMetSer 359

RESULT 4
US-09-764-864-1532
; Sequence 1532, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1532
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (134)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1532

Alignment Scores:
Pred. No.: 3,52e-101 Length: 250
Score: 116.00 Matches: 116
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.15% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-764-864-1532 (1-250)

QY 1251 GCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGT 1310
DB 135 AlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPhePheGly 154
QY 1311 CCCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1370
DB 155 ProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArg 174
QY 1371 GTCATTGGCAAGGTGCAAAACGGTGAACGAGTTGCAAGATTTCGAGGCGCTGAGGTG 1430

```

Db 175 ValileGlyLysGlyGlyThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 194
QY 1431 GTAGTACCAAGAGACCCCTGTAGAGACGACGAGTCTCGTGAATCATCGGA 1490
Db 195 ValValProArgaspGlnThrProaspGluAsnAspGlnValleValysIleGly 214
QY 1491 CATTTCTATCCAGTACAGTGGCTCAACGGAAGATCCGACATCTCTGCGCCCAAGTTAAG 1550
Db 215 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAlaGlnValLys 234
QY 1551 CAGCAGCATCAGAGGACAGAGTACACCGCCCGACGAGGAG 1598
Db 235 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgLys 250

RESULT 5
US-09-873-637-18
; Sequence 18, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-18

Alignment Scores:
Pred. No.: 3,54e-37 Length: 49
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.67% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-873-637-18 (1-49)
QY 714 ATCTGCCCCATAATACTTTAGGCGCTCTATTGGCAAGAGACGGAACCTGAAG 773
Db 1 IleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 20
QY 774 AAGGTAGACCAAGATACCGAGACAAAATCAACCATCTCTCGTTGCAAGACCTTACCCTT 833
Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40
QY 834 TACAACCCCTGAGAGACCATCACTGTG 860
Db 41 TyrAsnProGluArgThrIleThrVal 49

RESULT 6
US-09-873-637-20
; Sequence 20, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus

```

```

US-09-873-637-20
Alignment Scores:
Pred. No.: 3,19e-36 Length: 48
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.51% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-873-637-20 (1-48)
QY 1341 ATACGTGTGCCAGCATCAGCAGCTGCGCGGTCAATTCGCAAGTGGAAAAACGGTCAAC 1400
Db 1 IleArgValProAlaSerAlaAlaGlyValIleGlyLysGlyGlyThrValAsn 20
QY 1401 CAGTTGCAGAAATTTGACGGCAGCTGAGTGTGTAGTACCAAGACACCAAGACCCCTGATGAG 1460
Db 21 GluLeuGlnAsnLeuThrAlaAlaGluValValProArgaspGlnThrProaspGlu 40
QY 1461 AACGACACAGTCTCATCTCGTGAATC 1484
Db 41 AsnAspGlnValIleValLysIle 48

RESULT 7
US-09-873-637-17
; Sequence 17, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-17

Alignment Scores:
Pred. No.: 2,89e-35 Length: 47
Score: 47.00 Matches: 47
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.36% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-873-637-17 (1-47)
QY 471 CTCCTGTGCCCCACCCAGTATGTGGTGCATTTGGCAAGAGGGGGCCACCATCCGC 530
Db 1 LeuLeuValProThrGlnThrValGlyAlaIleIleGlyLysGlyAlaThrIleArg 20
QY 531 AACATCACAACCAAGACCCAGTCCAGATAGAGCTGCATAGGAGGAGAACCGAGGTGCA 590
Db 21 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 40
QY 591 GCTGAAAAGCCATCAGTGTG 611
Db 41 AlaGluLysAlaIleSerVal 47

RESULT 8
US-09-873-637-19
; Sequence 19, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

```

```
FILE REFERENCE: 960296.95131
CURRENT APPLICATION NUMBER: US/09/873,637
CURRENT FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 47
TYPE: PRT
ORGANISM: Mus musculus
US-09-873-637-19

Alignment Scores:
Pred. No.: 2,33e-33 Length: 47
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 45
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.04% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-873-637-19 (1-47)
QY 1095 GTGTTATCCGCCCGCCAGCAGTGGCGCATCATCGGCAAGAGGGGCGACATCAAA 1154
Db 1 ValPheilleProAlaGlnAlaValGlyAlaIlelleGlyLysGlyGlnHisIleLys 20
QY 1155 CAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGCACCAACCGAAACACCTGACTCC 1214
Db 21 GlnLeuSerArgPheAlaSerAlaSerIlelleLysIleAlaProGluThrProAspSer 40
QY 1215 AAGTTCGTATGTT 1229
Db 41 LysValArgMetVal 45

RESULT 9
US-10-117-982-472
; Sequence 472, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-472

Alignment Scores:
Pred. No.: 8.59e-25 Length: 70
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x US-10-117-982-472 (1-70)
QY 492 GTGGTGCCATTATTGGCAAGGAGGGGCCACCATCGCAACATCACAACAGACCCAG 551
```

```
Db 18 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgSnlleThrLysGlnThrGln 37
QY 552 TCCAAGATAGACGTGCATAGGAAGGAGAGCGACGTGCAGCTGAAAAA 599
Db 38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 10
US-10-313-986-472
; Sequence 472, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-472

Alignment Scores:
Pred. No.: 8.59e-25 Length: 70
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x US-10-313-986-472 (1-70)
QY 492 GTGGTGCCATTATTGGCAAGGAGGGGCCACCATCGCAACATCACAACAGACCCAG 551
Db 18 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgSnlleThrLysGlnThrGln 37
QY 552 TCCAAGATAGACGTGCATAGGAAGGAGAGCGACGTGCAGCTGAAAAA 599
Db 38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 11
US-10-408-765A-2088
; Sequence 2088, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2088
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2088
```

Alignment Scores:
 Pred. No.: 6.99e-25 Length: 422
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 16 Gaps: 0

US-09-270-437D-7 (1-1946) x US-10-408-765A-2088 (1-422)

QY 492 GTGGTGCCATTATTGCGAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 50 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 69
 QY 552 TCCAAGATAGACGTGCATAGGAAGGAGAACGCGAGGTGCAGCTGAAAAA 599
 Db 70 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 85

RESULT 12

US-09-735-705-176
 ; Sequence 176, Application US/09735705
 ; Patent No. US20020052329A1

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
 APPLICANT: Fan, Liqun
 APPLICANT: Kalos, Michael D.
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Hosken, Nancy
 APPLICANT: Fanger, Gary R.
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Henderson, Robert A.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Fanger, Neil
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.455C14
 CURRENT APPLICATION NUMBER: US/09/735,705
 CURRENT FILING DATE: 2000-12-12
 NUMBER OF SEQ ID NOS: 419
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 176
 LENGTH: 579
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-735-705-176

Alignment Scores:
 Pred. No.: 6.74e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-735-705-176 (1-579)

QY 492 GTGGTGCCATTATTGCGAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAAGATAGACGTGCATAGGAAGGAGAACGCGAGGTGCAGCTGAAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 13

US-09-735-705-348
 ; Sequence 348, Application US/09735705
 ; Patent No. US20020052329A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Tongtong
 APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Hosken, Nancy
 APPLICANT: Fanger, Gary R.
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Henderson, Robert A.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Fanger, Neil
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.455C14
 CURRENT APPLICATION NUMBER: US/09/735,705
 CURRENT FILING DATE: 2000-12-12
 NUMBER OF SEQ ID NOS: 419
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 348
 LENGTH: 579
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-735-705-348

Alignment Scores:
 Pred. No.: 6.74e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-735-705-348 (1-579)

QY 492 GTGGTGCCATTATTGCGAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAAGATAGACGTGCATAGGAAGGAGAACGCGAGGTGCAGCTGAAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 14

US-09-850-716A-176
 ; Sequence 176, Application US/09850716A
 ; Patent No. US20020115139A1

GENERAL INFORMATION:

APPLICANT: Kalos, Michael D.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Retter, Marc W.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.455C15
 CURRENT APPLICATION NUMBER: US/09/850,716A
 CURRENT FILING DATE: 2001-05-07
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 176
 LENGTH: 579
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-850-716A-176

Alignment Scores:
 Pred. No.: 6.74e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: 9 Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-850-716A-176 (1-579)

QY 492 GTGGTGCCATTATTGCGAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551

Db 207 ValGlyAlaIleGlyGlyGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAGATAGACGTGCATAGGAGGAGACGAGCGAGTGCAGCTGAAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 15

US-09-850-716A-348
 ; Sequence 348, Application US/09850716A
 ; Patent No. US20020115139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Retter, Marc W.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C15
 ; CURRENT APPLICATION NUMBER: US/09/850,716A
 ; CURRENT FILING DATE: 2001-05-07
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-850-716A-348

Alignment Scores:
 Pred. No.: 6,74e-25 Length: 579
 Score: 36.00 Matches: 36
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.63% Indels: 0
 DB: Gaps: 0

US-09-270-437D-7 (1-1946) x US-09-850-716A-348 (1-579)

QY 492 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCACAAACAGACCCAG 551
 Db 207 ValGlyAlaIleGlyGlyGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
 QY 552 TCCAGATAGACGTGCATAGGAGGAGACGAGCGAGTGCAGCTGAAAAA 599
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

Search completed: August 6, 2004, 14:13:17
 Job time : 100.225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 29.5863 Seconds
(without alignments)
12653.723 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 639
Sequence: 1 gctgtagggagggtggg.....atttccttcaggttttaaaa 1946

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool/p/US03270437/runat_06082004_141221_1829/app_query.fasta_1.5582
-DB=PIR_78 -QFMT=fastan -SUFFIX=oligo806.rpr -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=829 -NCPU=6 -ICPU=3
-USER=US03270437@cgn_1_133@runat_06082004_141221_1829 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	1.3	95	H95092	hypothetical prote
C 2	8	1.3	95	E97960	hypothetical prote
C 3	8	1.3	150	A87256	hypothetical prote
C 4	8	1.3	185	I39736	hypothetical prote
C 5	8	1.3	217	AD2683	Conserved hypotet
C 6	8	1.3	221	AF3195	transcription regu
C 7	8	1.3	223	A83859	menaquinol-cytochr
C 8	8	1.3	224	I39943	menaquinol-cytochr
C 9	8	1.3	227	I470036	capsular polysacch
C 10	8	1.3	242	F75433	probable phosphoes
C 11	8	1.3	247	B90733	probable major tai
C 12	8	1.3	249	D85583	probable tail comp
C 13	8	1.3	253	B97465	hypothetical prote
C 14	8	1.3	259	G64831	probable membrane

ALIGNMENTS

RESULT 1

H95092

hypothetical protein SP0800 [imported] - streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C/Accession: H95092

R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heic

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A/Reference number: A95000; PMID:21357209; PMID:11463916

A/Accession: H95092

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-95 <KUR>

A/Cross-references: GB:AE005672; PIDN:AAK74937.1; PID:gl4972276; GSPDB:GN00164; TIGR:SP.

A/Experimental source: strain TIGR4

C/Genetics:

A/Gene: SP0800

Alignment Scores:			
Pred. No.:	76	Length:	95
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.26%	Indels:	0
DB:	2	Gaps:	0

US-09-270-437D-7 (1-1946) x H95092 (1-95)

Qy 1940 AACCTGAAGGAATTTTATTTTC 1917

Db 36 AsnLeuIysGlulLeuPhePhe 43

RESULT 2

E97960

hypothetical prote
hypothetical prote
heterodisulfide re
agmatine ureohydro
hypothetical prote
fibr protein (AF19
hypothetical prote
conserved hypotet
hypothetical prote
probable aminotran
hypothetical prote
G box-binding prot
G box-binding prot
neuron-specific si
hypothetical prote
hypothetical prote
probable AMP nucle
structural protein
splicing factor-li
F3F19.21 protein -
probable ATP-depen
anthranilate synth
asparagine synthas
chromogranin B pre
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable integral
hypothetical prote

C 15 8 1.3 259 2 C90754
C 16 8 1.3 259 2 A85618
C 17 8 1.3 301 2 AF2223
C 18 8 1.3 306 2 AF0876
C 19 8 1.3 340 2 AB2584
C 20 8 1.3 340 2 B97366
C 21 8 1.3 346 2 T46916
C 22 8 1.3 351 2 AC3090
C 23 8 1.3 351 2 G98196
C 24 8 1.3 357 2 B81396
C 25 8 1.3 367 2 T36278
C 26 8 1.3 374 2 T07887
C 27 8 1.3 376 2 T07882
C 28 8 1.3 403 2 JC5269
C 29 8 1.3 457 2 B87269
C 30 8 1.3 475 2 G70861
C 31 8 1.3 532 2 E95949
C 32 8 1.3 546 2 AI0278
C 33 8 1.3 555 2 T30349
C 34 8 1.3 573 2 C85433
C 35 8 1.3 573 2 C86266
C 36 8 1.3 600 2 T09676
C 37 8 1.3 638 2 F75547
C 38 8 1.3 645 2 H96011
C 39 8 1.3 646 1 S15901
C 40 8 1.3 680 2 T25832
C 41 8 1.3 731 2 T04855
C 42 8 1.3 735 2 A84829
C 43 8 1.3 755 2 T47731
C 44 8 1.3 759 2 T39090
C 45 8 1.3 812 2 T34180

hypothetical protein spr0709 [imported] - Streptococcus pneumoniae (strain R6)
 C;Species: Streptococcus pneumoniae
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C;Accession: E97960
 R;Hoskins J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; BURGERT, S.; DeHoff, B.S.; E
 y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 e, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: E97960
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-95 <KUR>
 A;Cross-references: GB:AE007317; PIDN:AAK99513.1; PID:gl5458299; GSPDB:GN00174
 C;Genetics:
 A;Gene: spr0709

Alignment Scores:
 Pred. No.: 76 Length: 95
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.26% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x E97960 (1-95)

QY 1940 AACCTGAAGGAATTTTATTTTC 1917
 |||||
 Db 36 AsnLeuLysGluIleLeuPhePhe 43

RESULT 3
 A87256
 hypothetical protein CC0058 [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C;Accession: A87256
 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 n, J.; Laub, M.R.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: A87256
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-150 <STO>
 A;Cross-references: GB:AE005673; NID:gl13421153; PIDN:AAK22045.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC0058
 C;Superfamily: Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimI

Alignment Scores:
 Pred. No.: 71.6 Length: 150
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.25% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x A87256 (1-150)

QY 8 CGAGAGGGCTGGGGGCTGCTCTG 31
 |||||
 Db 79 ArgArgGlyTTPGlyAlaAlaLeu 86

RESULT 4
 I39736
 hypothetical protein 6 - Anabaena variabilis
 C;Species: Anabaena variabilis
 C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999

C;Accession: S68185; I39736
 R;Schmitz, O.; Boison, G.; Hilscher, R.; Hundeshagen, B.; Zimmer, W.; Lottspeich, F.; Bo
 Eur. J. Biochem. 233, 266-276, 1995
 A;Title: Molecular biological analysis of a bidirectional hydrogenase from cyanobacteria
 A;Reference number: I39730; MUID:96061958; PMID:7588754
 A;Accession: S68185
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-185 <SCH>
 A;Cross-references: EMBL:X79285; NID:gl032475; PIDN:CAA55879.1; PID:gl032482
 A;Experimental source: ATCC 29413
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994

Alignment Scores:
 Pred. No.: 69.6 Length: 185
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.26% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x I39736 (1-185)

QY 716 GATCTTCAGGGGAACCTCGTCGCG 693
 |||||
 Db 90 AsnLeuGlnGlyAsnLeuValSer 97

RESULT 5
 AD2683
 conserved hypothetical protein Atu0868 [imported] - Agrobacterium tumefaciens (strain C5
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C;Accession: AD2683
 R;Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AD2683
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-217 <KUR>
 A;Cross-references: GB:AE008688; PIDN:AAI41882.1; PID:gl7739245; GSPDB:GN00186
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu0868
 A;Map position: circular chromosome

Alignment Scores:
 Pred. No.: 68.2 Length: 217
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.26% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x AD2683 (1-217)

QY 1899 TCCACCTCTTGGTGGGCTCGA 1876
 |||||
 Db 103 SerThrLeuValGlyProGly 110

RESULT 6
 AF3195
 transcription regulator, TetR family Atu5291 [imported] - Agrobacterium tumefaciens (str
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C;Accession: AF3195
 R;Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

A:Accession: I39943
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-224 <RES>
A:Cross-references: GB:D45410; NID:G902924; PID:G643677
C:Genetics:
A:Gene: petB
A:Start codon: GTG
C:Superfamily: cytochrome b6; cytochrome b6 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase
F:16-224/Domain: cytochrome b6 homology <b6>
F:43/Binding site: heme (Cys) (covalent) (probably high potential) #status predicted
F:94,196/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:108,211/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Alignment Scores:
Pred. No.: 67.9 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x I39943 (1-224)

QY 1024 CTCACGACGCGTACTGGCGTG 1047
|||||
Db 200 LeuProAlaLeuLeuGlyLeu 207
|||||

RESULT 9
A70036
capsular polysaccharide biosynthesis homolog yveL - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: A70036
A:Authors: R.; Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Etian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
ioeth, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.H.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Sefforge, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A70036
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-227 <KUN>
A:Cross-references: GB:J299121; GB:AL009126; NID:G2635827; PIDN:CAB15441.1; PID:G2635949
A:Experimental source: strain 168
C:Genetics:
A:Gene: yveL
C:Superfamily: capsular polysaccharide biosynthesis protein cpsC

Alignment Scores:
Pred. No.: 67.8 Length: 227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x A70036 (1-227)

QY 1605 CTGGTACATCTCCGTCGCTGGG 1582
|||||

```

Db      48 LeuValThrSerSerValProGly 55
RESULT 10
F75433
probable phosphoesterase (EC 3.1.1.-) D1119 [similarity] - Deinococcus radiodurans (str
C;Species: Deinococcus radiodurans
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: F75433
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: F75433
A;Molecule type: DNA
A;Residues: 1-242 <WHI>
A;Cross-references: GB:AE0001962; GB:AE000513; NID:G6458955; PIDN:AAF0693.1; PID:G645895
A;Experimental source: strain R1
C;Comment: This sequence has motifs characteristic of a variety of phosphoesterases.
C;Genetics:
A;Gene: D1119
A;Map position: 1
A;Superfamily: Chlamydia trachomatis probable phosphoesterase CT488; phosphoesterase cor
C;Keywords: hydrolase
F;2-84/Domain: phosphoesterase core homology <PEC>
Alignment Scores:
Pred. No.: 67.2 Length: 242
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 1 Gaps: 0
US-09-270-437D-7 (1-1946) x F75433 (1-242)
QY 115 CGCTCTTACCGGAGACTGAG 92
Db 136 ArgLeuLeuThrArgGluAlaGlu 143
RESULT 11
B90733
probable major tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: B90733
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gawawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA834257.1; PID:G13360293; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: Ecs0834
C;Superfamily: phage lambda major tail protein V
Alignment Scores:
Pred. No.: 67.1 Length: 247
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 2 Gaps: 0
US-09-270-437D-7 (1-1946) x B90733 (1-247)
QY 597 AAAGCATCAGTGTGCCTCCACC 620
Db 193 LysAlaIleSerValHisSerThr 200
RESULT 12
D85583
probable tail component of prophage CP-933K Z0972 [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D85583
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85583
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: GB:AE005174; NID:G12513740; PIDN:AAG55128.1; GSPDB:GN00145; UWGP:209
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z0972
C;Superfamily: phage lambda major tail protein V
Alignment Scores:
Pred. No.: 67 Length: 249
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 2 Gaps: 0
US-09-270-437D-7 (1-1946) x D85583 (1-249)
QY 597 AAAGCATCAGTGTGCCTCCACC 620
Db 195 LysAlaIleSerValHisSerThr 202
RESULT 13
B97465
hypothetical protein AGR_C1585 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: B97465
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97465
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-253 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86675.1; PID:G15155859; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C1585
A;Map position: circular chromosome
Alignment Scores:
Pred. No.: 66.9 Length: 253
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 2 Gaps: 0
US-09-270-437D-7 (1-1946) x B97465 (1-253)
QY 1899 TCACCTCTTGTGGGCTGGA 1876
Db 139 SerThrLeuValGlyProGly 146
RESULT 14

```

Search completed: August 6, 2004, 13:33:27
Job time : 33.5863 secs

G64831
probable membrane protein ybcC - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: G64831
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64831
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-259 <BLAT>
A:Cross-references: GB:AE000194; GB:U00096; NID:gl787148; PIDN:AAC74006.1; PID:gl787150;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ybcC
C:Superfamily: conserved hypothetical protein aq_1986
C:Keywords: transmembrane protein
F:13-29/Domain: transmembrane #status predicted <TM1>
F:39-55/Domain: transmembrane #status predicted <TM2>

Alignment Scores:
Pred. No.: 66.6 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x G64831 (1-259)
QY 717 GGATCTTCAGGGGACCTCGTCAG 694
Db 251 GlySerSerGlyGluProArgGln 258
|||||
|||||

RESULT 15
C90754
hypothetical protein ECs1003 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C90754
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A93629; MUID:21156231; PMID:11258796
A:Reference number: A93629; MUID:21156231; PMID:11258796
A:Accession: C90754
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA034426.1; PID:gl3360462; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs1003
C:Superfamily: conserved hypothetical protein aq_1986

Alignment Scores:
Pred. No.: 66.6 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-7 (1-1946) x C90754 (1-259)
QY 717 GGATCTTCAGGGGACCTCGTCAG 694
Db 251 GlySerSerGlyGluProArgGln 258
|||||
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 17.1191 Seconds
(without alignments)
11838.037 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 639
Sequence: 1 gctgtagcggagggtggg.....atttccttcagggttttaaaa 1946

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1
Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US09270437/runat_06082004_141220_1803/app_query.fasta_1.5582
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=oligo806.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437@cgn_1.57/runat_06082004_141220_1803 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	1.3	166	1 ING_CAMBA	Q865w6 camelus bac
C 2	8	1.3	166	1 ING_CAMBA	Q865x1 lama glama
C 3	8	1.3	224	1 OCRE_BACTC	Q45658 bacillus th
C 4	8	1.3	227	1 YVEL_BACSU	P71051 bacillus su
C 5	8	1.3	229	1 YVBC_ECOLI	P36565 escherichia
C 6	8	1.3	267	1 RS2_DROME	P31009 drosophila
C 7	8	1.3	371	1 GP85_BRARE	Q91919 brachydanio
C 8	8	1.3	403	1 STAC_MOUSE	P97306 mus musculus
C 9	8	1.3	473	1 PHDK_NOCSK	O24723 nocardioides
C 10	8	1.3	512	1 DFN5_MOUSE	Q922d3 mus musculus
C 11	8	1.3	600	1 Y018_MYCGE	P47264 mycoplasma
C 12	8	1.3	622	1 3BP1_HUMAN	Q9Y1l3 homo sapien
C 13	8	1.3	646	1 SGI_EOVIN	P23389 bos taurus
C 14	8	1.3	739	1 BAC1_MOUSE	P97302 mus musculus
C 15	8	1.3	815	1 GYRE_MYXXA	O33367 myxococcus
C 16	8	1.3	1030	1 Y018_MYCPN	P75093 mycoplasma
C 17	8	1.3	1473	1 TOP2_ARATH	P30182 arabidopsis
C 18	8	1.3	1648	1 YJ9H_YEAST	P47171 saccharomyc

C 19	8	1.3	1692	1 RIM1_HUMAN	Q86ur5 homo sapien
C 20	8	1.3	2567	1 M18B_HUMAN	Q8iug5 homo sapien
C 21	8	1.3	3176	1 CA36_HUMAN	P12111 homo sapien
C 22	8	1.3	4451	1 GRSB_BACBR	P14688 b gramicidin
C 23	8	1.3	5035	1 RYR1_PIG	P16960 sus scrofa
C 24	8	1.3	5037	1 RYR1_RABIT	P11716 oryctolagus
C 25	8	1.3	5038	1 RYR1_HUMAN	P21817 homo sapien
C 26	7	1.1	85	1 RL32_VIBCH	Q9kqk3 vibrio chol
C 27	7	1.1	85	1 RL32_VIBCH	Q9kqk3 vibrio chol
C 28	7	1.1	98	1 FIS_ERWCA	PIS33 YEAST
C 29	7	1.1	98	1 FIS_ERWCA	Q82ax8 erwina car
C 30	7	1.1	101	1 CBGR_CLOAB	P24132 clostridium
C 31	7	1.1	110	1 LV2J_HUMAN	P01713 homo sapien
C 32	7	1.1	113	1 GVK1_HALN1	P24375 halobacteri
C 33	7	1.1	114	1 RL42_BRUNA	P24375 halobacteri
C 34	7	1.1	124	1 VA31_VACCC	P21096 vaccinia vi
C 35	7	1.1	124	1 VA31_VACCV	P24760 vaccinia vi
C 36	7	1.1	136	1 C17_HUMAN	Q9nrri homo sapien
C 37	7	1.1	136	1 FLIS_VIBCH	Q9kqk5 vibrio chol
C 38	7	1.1	140	1 VA31_VARV	P33848 variola vir
C 39	7	1.1	143	1 GLPA_PANTR	Q28913 pan troglod
C 40	7	1.1	150	1 ME15_SCHPO	Q96w86 schizosacch
C 41	7	1.1	153	1 HS11_SOYBN	P02519 glycine max
C 42	7	1.1	164	1 ING_CHICK	P49708 gallus gall
C 43	7	1.1	164	1 ING_COTJA	O57571 coturnix co
C 44	7	1.1	164	1 ING_MELGA	O57603 meleagris g
C 45	7	1.1	164	1 ING_NUMME	O73915 numida mele

ALIGNMENTS

RESULT 1

ING_CAMBA STANDARD; PRT; 166 AA.
AC Q865W6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Interferon gamma precursor (IFN-gamma).
GN IFNG.
OS Camelus bactrianus (Bactrian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9837;
RN [1]
RP SEQUENCE FROM N.A.
RA Raadan O., Lee S., Yoshida R., Chang K., Chashi K., Sugimoto C.,
RA Onuma M.;
RT "Cloning and sequence analysis of cytokine cDNAs of llama and camel.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Produced by lymphocytes activated by specific antigens
CC or mitogens. IFN-gamma, in addition to having antiviral activity,
CC has important immunoregulatory functions. It is a potent activator
CC of macrophages, it has antiproliferative effects on transformed
CC cells and it can potentiate the antiviral and antitumor effects of
CC the type I interferons (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: Released primarily from activated T
CC lymphocytes.
CC -!- SIMILARITY: Belongs to the type II (or gamma) interferon family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AB107657; BAC75394.1; -
DR InterPro; IPR002069; IFN-gamma.
DR Pfam; PF00714; IFN-gamma; 1.
DR

DR ProDom; PD002435; IFN-gamma; 1.
KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 166 INTERFERON GAMMA.
FT CARBOHYD 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 166 AA; 19402 MW; 4547EC4FFC93655 CRC64;

Alignment Scores:

Pred. No.: 37.4 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x ING_CAMBA (1-166)

Qy 536 GATGTCGGATGGTGGCCCCCTC 513

Db 44 AspValAlaAspGlyGlyProLeu 51

RESULT 2

ID_ING_LAMGL

AC Q865X1; STANDARD; PRT; 166 AA.

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Interferon gamma precursor (IFN-gamma).

GN IFNG.

OS Lama glama (Llama).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.

OX NCBI_TaxID=9844;

RN [1]

RP SEQUENCE FROM N.A.

RA Raanan O., Lee S., Yoshida R., Chang K., Ohashi K., Sugimoto C.,

RA Onuma M.; and sequence analysis of cytokine cDNAs of llama and camel.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Produced by lymphocytes activated by specific antigens

CC or mitogens. IFN-gamma, in addition to having antiviral activity,

CC has important immunoregulatory functions. It is a potent activator

CC of macrophages, it has antiproliferative effects on transformed

CC cells and it can potentiate the antiviral and antitumor effects of

CC the type I interferons (By similarity).

CC -1- SUBUNIT: Homodimer (By similarity).

CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -1- TISSUE SPECIFICITY: Released primarily from activated T

CC lymphocytes.

CC -1- SIMILARITY: Belongs to the type II (or gamma) interferon family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AB107652; BAC75389.1; -

CC InterPro; IPR002069; IFN-gamma; 1.

CC Pfam; PF00714; IFN-gamma; 1.

CC ProDom; PD002435; IFN-gamma; 1.

KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.

FT SIGNAL 1 20 BY SIMILARITY

FT CHAIN 21 166 INTERFERON GAMMA.

FT CARBOHYD 39 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 106 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 166 AA; 19475 MW; 23937B814759328F CRC64;

Alignment Scores:

Pred. No.: 37.4 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x ING_LAMGL (1-166)

Qy 536 GATGTCGGATGGTGGCCCCCTC 513

Db 44 AspValAlaAspGlyGlyProLeu 51

RESULT 3

ID_QCRB_BACTC

AC Q45658; STANDARD; PRT; 224 AA.

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Menquinol-cytochrome c reductase cytochrome B subunit.

GN QCRB.

OS Bacillus thermodenitrificans.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.

OX NCBI_TaxID=33940;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K1041;

RX MEDLINE=96218169; PubMed=8647852;

RA Some N., Tsuchiya N., Inoue M., Noguchi S.;

RT "Bacillus stearotherophilus qcr operon encoding Rieske FeS protein, c

RT cytochrome b6, and a novel-type cytochrome c1 of quinol-cytochrome c

RT reductase";

RT J. Biol. Chem. 271:12457-12462(1996).

CC -1- FUNCTION: Component of the menquinol-cytochrome c reductase

CC complex.

CC -1- COFACTOR: Two heme groups which are not covalently bound to the

CC protein (By similarity).

CC -1- SUBUNIT: THE MAIN SUBUNITS OF THE MENAQUINONE-CYTOCHROME C COMPLEX

CC ARE: CYTOCHROME B, THE RIESKE PROTEIN AND A 22/29 KDa CYTOCHROME

CC B/C SUBUNIT.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY. CORRESPONDS TO THE

CC AMINO END OF MITOCHONDRIAL CYTOCHROME B.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D83789; BAA12117.1; -

CC InterPro; IPR005797; Cytochrome_b; 1.

CC Pfam; PF00033; Cytochrome_b; 1.

CC PROSITE; PS00192; CYTOCHROME_B_HEME; 1.

KW Electron transport; Heme; Transmembrane.

FT TRANSMEM 37 57 POTENTIAL.

FT TRANSMEM 96 116 POTENTIAL.

FT TRANSMEM 126 146 POTENTIAL.

FT TRANSMEM 195 215 POTENTIAL.

FT METAL 94 94 POTENTIAL.

FT METAL 108 108 IRON (HEME 1 AXIAL LIGAND)

FT METAL 196 196 (BY SIMILARITY).

FT METAL 211 211 IRON (HEME 2 AXIAL LIGAND)

FT METAL 211 211 (BY SIMILARITY).

FT METAL 211 211 IRON (HEME 1 AXIAL LIGAND)

FT METAL 211 211 (BY SIMILARITY).

FT METAL 211 211 IRON (HEME 2 AXIAL LIGAND)

FT METAL 211 211 (BY SIMILARITY).

SQ SEQUENCE 224 AA; 25411 MW; AD0581P4E7F0518 CRC64;

Alignment Scores:


```

Pred. No.: 36.3 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x QCRB_BACTC (1-224)
QY 1024 CTCGCCAGCAGCTTACTGGGGTGG 1047
DB 200 LeuProAlaLeuLeuGlyLeu 207

RESULT 4
YVEL_BACSU
ID YVEL_BACSU STANDARD; PRT; 227 AA.
AC P71051; O08170;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Putative tyrosine-protein kinase yvel (EC 2.7.1.112).
GN YVEL OR BS034360.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124193; PubMed=8969506;
RA Fabret C., Quentin Y., Chapal N., Guiseppi A., Haiech J., Denizot F.;
"Integrated mapping and sequencing of a 115 kb DNA fragment from
RT Bacillus subtilis: sequence analysis of a 21 kb segment containing
RT the sigL locus.";
RL Microbiology 142:3089-3096(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Denizot F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borcherst S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golithly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpsira P., Tognoni A.,
RA Tosa V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weizenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";

```

```

RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SIMILARITY: Belongs to the cpsD/capB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z71928; CAA96490.1; -
DR EMBL; Z94043; CAB08024.1; -
DR EMBL; Z99121; CAB15441.1; -
DR PIR; A70036; A70036.
DR Subtilist; BG11860; Yvel.
DR InterPro; IPR005702; EPS_synthesis.
DR TIGRFAMs; TIGR01007; eps_fam; 1.
KW Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW Complete proteome.
SQ SEQUENCE 227 AA; 24674 MW; 81C8B9D75278FFA9 CRC64;

Alignment Scores:
Pred. NO.: 36.2 Length: 227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x YVEL_BACSU (1-227)
QY 1605 CTGGTCACCTTCCTCCGCTGGG 1582
DB 48 LeuValThrSerValProGly 55

RESULT 5
YCBC_ECOLI
ID YCBC_ECOLI STANDARD; PRT; 259 AA.
AC P36565; P75846;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycbc.
GN YCBC OR B0920 OR Z1267 OR ECS1003.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X69120; CAA49872.1; --
DR EMBL; AE003626; AAF52822.1; --
DR EMBL; U01334; AAC34198.1; --
DR EMBL; U01335; AAA87053.1; --
DR PIR; S30395; S30395.
DR HSP; P02357; 1PKP.
DR FlyBase; FBgn004867; sop.
DR InterPro; IPR000851; Ribosomal_S5.
DR InterPro; IPR005324; Ribosomal_S5_C.
DR InterPro; IPR005711; Ribosomal_S5_e/a.
DR Pfam; PF00333; Ribosomal_S5_C; 1.
DR Pfam; PF03719; Ribosomal_S5_C; 1.
DR TIGRfam; TIGR01020; rpsE arch; 1.
DR PROSITE; PS00585; RIBOSOMAL_S5; 1.
DR PROSITE; PS00881; S5_DSRED; 1.
KW Ribosomal protein.
FT DOMAIN 85 148 S5 DREM.
FT INTERPRO 19 20 GG -> PP (IN REF. 3).
FT CONFLICT 194 194 K -> R (IN REF. 3).
SQ SEQUENCE 267 AA; 28899 MW; ADA22CD28F100743 CRC64;

Alignment Scores:
Pred. No.: 35.6 Length: 267
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x RS2_DROME (1-267)
QY 375 CGAGGGGGCTTTCGCTCTCGGGT 398
Db 12 ArgGlyGlyPheGlySerArgGly 19

RESULT 7
GP85_BRARE
ID GP85_BRARE STANDARD; PRT; 371 AA.
AC Q91919;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Super conserved receptor expressed in brain 2.
GN SREB2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20294882; PubMed=10833454;
RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuichi K.,
RT "An evolutionarily conserved G-protein coupled receptor family, SREB,
RT expressed in the central nervous system."
RL Biochem. Biophys. Res. Commun. 272:576-582 (2000).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- FUNCTION: Orphan receptor.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB040805; BAA96651.1; --
DR ZFIN; ZDB-GENE-000710-2; sre2.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G PROTEIN RECP FL 1; FALSE_NEG.
DR PROSITE; PS00262; G-PROTEIN RECP FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Multigene family.
FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 27 47 1 (POTENTIAL).
FT DOMAIN 48 58 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 59 79 2 (POTENTIAL).
FT DOMAIN 80 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 117 3 (POTENTIAL).
FT DOMAIN 118 138 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 5 (POTENTIAL).
FT DOMAIN 211 287 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 288 308 6 (POTENTIAL).
FT DOMAIN 309 321 7 (POTENTIAL).
FT TRANSMEM 322 342 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 343 371 BY SIMILARITY.
FT DISULFID 95 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 371 AA; 41954 MW; F6F6175ED3A348C2 CRC64;

Alignment Scores:
Pred. No.: 34.5 Length: 371
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.28% Indels: 0
DB: Gaps: 0

US-09-270-437D-7 (1-1946) x GP85_BRARE (1-371)
QY 1040 AGTACGCTGCTGGGAGCGCGG 1017
Db 264 SerAsnAlaAlaGlyArgArgArg 271

RESULT 8
STAC_MOUSE
ID STAC_MOUSE STANDARD; PRT; 403 AA.
AC P97306;
DT 16-Oct-2001 (Rel. 40, Created)
DT 16-Oct-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Stac protein (SRC homology 3 and cysteine-rich domain protein).
GN STAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C3H; TISSUE=Brain;
RX MEDLINE=97115677; PubMed=8954993;
RA Suzuki H., Kawai J., Taga C., Yaoi T., Hara A., Hirose K.,
RA Hayashizaki Y., Watanabe S.,
RT "Stac, a novel neuron-specific protein with cysteine-rich and SH3 domains."
RL Biochem. Biophys. Res. Commun. 225:902-909 (1996).
CC -!- FUNCTION: Probably involved in a neuron-specific signal transduction.


```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Cochlea;
RX MEDLINE=98442658; PubMed=9771715;
RA Van Laer L., Huizinga E.H., Verstreken M., van Zuijlen D.,
RA Wauters J.G., Bessuyt P.J., Van de Heyning P., McGuire W.T.,
RA Smith R.J.H., Willems P.J., Legan P.K., Richardson G.P., Van Camp G.;
RT "Nonsyndromic hearing impairment is associated with a mutation in
RT DFNAS5";
RL Mat. Genet. 20:194-197(1998).
CC -!- SIMILARITY: BELONGS TO THE DFN5 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF073309; AAC69325.1; -
CC DR MGD; MGI:189850; Dfnas5.
CC DR InterPro; IPR007677; DFNAS5.
CC DR InterPro; IPR007681; MGBL.
CC PFam; PF04598; DFNAS5; 1.
CC SQ SEQUENCE 512 AA; 56630 MW; 13AFB8627773C4A5 CRC64;

Alignment Scores:
Pred. No.: 33.3 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x DFN5_MOUSE (1-512)
QY 854 GATGTCCTCTCAGGTTCTTAAG 831
DB 279 AspGlyProLeuArgValVallys 286

RESULT 11
Y018_MYCGE STANDARD; PRT; 600 AA.
AC Y018_MYCGE Q49302;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical helicase MG018.
GN MG018.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]_TaxID=2097;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Merrick J.M.,
RA Nguyen D.T., Dougherty B.A., Burt K.F., Hu P.-C., Lucier T.S.,
RA Tomb J.-F., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP REVISIONS.
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

```

```

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Burt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 209-309 AND 371-471 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Burt K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -!- MISCELLANEOUS: IN M.PNEUMONIAE, A SINGLE ORF SPANS M.GENITALIUM
CC MG016, MG017 AND MG018.
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 helicase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39681; AAC71234.1; -
CC DR EMBL; U02179; RAD12465.1; -
CC DR EMBL; U01757; RAD10571.1; -
CC PIR; T09676; T09676.
CC DR TIGR; MG018; -
CC DR InterPro; IPR001410; DEAD.
CC DR InterPro; IPR001650; Helicase_C.
CC DR InterPro; IPR000330; SNF2_N.
CC PFam; PF00271; helicase_C_1.
CC PFam; PF00176; SNF2_N; 1.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELICC; 1.
KW Hypothetical protein; Helicase; ATP-binding; Complete proteome.
FT NP_BIND 172 179 ATP (POTENTIAL).
FT SITE 271 274 DEAD BOX.
FT CONFLICT 462 462 D -> S (IN REF. 3).
FT SEQUENCE 600 AA; 68873 MW; 75A82D909F4A877A CRC64;

Alignment Scores:
Pred. No.: 32.8 Length: 600
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x Y018_MYCGE (1-600)
QY 1410 AATTTCAGCGCAGCTGAGTGGTA 1433
DB 515 AsnLeuThrAlaAlaGluValVal 522

RESULT 12
3BP1_HUMAN
ID 3BP1_HUMAN STANDARD; PRT; 622 AA.
AC Q9X1L3; O56H05;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE SH3-domain binding protein 1 (3BP-1).
GN SH3BP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;

```

RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Beggley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Karshaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.T.,
RA Martyn I.D., Mashregi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Yoshinaka S., Kawaaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shinohara A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis S., Bentley J.D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Gosla D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.,
RA "The DNA sequence of human chromosome 22.",
RL Nature 402:489-495(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Binds differentially to the SH3 domains of certain
proteins of signal transduction pathways. This protein binds
preferentially to c-abl proto-oncogene, SRC and GRB2. Shows
GAP activity for Rac-related proteins but not for Rho- or
Ras-related proteins. It inhibits PDGF-induced membrane ruffling
mediated by Rac (By similarity).
CC

CC -1- SIMILARITY: Contains 1 Rho-GAP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z83844; CAB42896.1; --
DR EMBL: BC008282; AA08282.1; --
DR HSSP: Q07960; IRGP.
DR Genew: HGNC:10824; SH3BP1.
DR InterPro: IPR008936; RhoGAP.
DR InterPro: IPR000198; RhoGAP.
DR Pfam: PF00620; RhoGAP; 1.
DR SMART: SM00324; RhoGAP; 1.
DR PROSITE: PS00336; RHO GAP; 1.
KW GTPase activation; SH3-binding.
FT DOMAIN 196 390 RHO-GAP.
FT SITE 538 546 SH3-BINDING.
FT DOMAIN 113 117 POLY-GLU.
FT CONFLICT 229 229 MISSING (IN REF. 2).
FT CONFLICT 486 492 MISSING (IN REF. 1).
SQ SEQUENCE 622 AA; 66764 MW; 1F52C758DC7B9C06 CRC64;

Alignment Scores:
Pred. No.: 32.7 Length: 622
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.28% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x 3BP1_HUMAN (1-622)
QY 28 AGCAGCCCCCAGCCCTCGCTA 5
DB 506 SerSerProProAlaProProLeu 513

RESULT 13
SG1_BOVIN STANDARD; PRT; 646 AA.
ID_SG1_BOVIN
AC P23389; O02707;
DT 01-NOV-1991 (Rel. 20, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Secretogranin I precursor (Sgi) (Chromogranin B) (Cgb) [Contains: GAWK
DE peptide; Secretolytin].
GN CHGB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal chromaffin;
RX MEDLINE=91223091; PubMed=2025642;
RA Bauer J.W., Fischer-Colbrie R.,
RT "Primary structure of bovine chromogranin B deduced from cDNA
sequence.",
RL Biochim. Biophys. Acta 1089:124-126(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal medulla;
RX MEDLINE=97282586; PubMed=9136897;
RA Yoo S.H., Kang Y.K.,
RT "Identification of the secretory vesicle membrane binding region of
chromogranin B.",
RL FEBS Lett. 406:259-262(1997).
RN [3]

RP SEQUENCE OF 21-646 FROM N.A.
RC TISSUE=Adrenal medulla;
RA Grandy D.K., Leduc R., Makam H., Flanagan T., Dilberto E.J.,
RA Thomas G., Civelli O., Viveros O.H.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 634-646.
RC TISSUE=Adrenal chromaffin;
RA MEDLINE=95262699; PubMed=7744058;
RA Strub J.-M., Garcia-Sabione P., Lanning K., Taupenot L., Hubert P.,
RA van Dorsselaer A., Anis D., Metz-Boutigue M.-H.;
RT "Processing of chromogranin B in bovine adrenal medulla.
RT Identification of secretory, the endogenous C-terminal fragment of
RT residues 614-626 with antibacterial activity.";
RL Eur. J. Biochem. 229:358-368 (1995).
RN [5]
RP CHARACTERIZATION OF SECRETOLYTIN.
RC MEDLINE=96184581; PubMed=8603705;
RA Strub J.M., Hubert P., Nullans G., Anis D., Metz-Boutigue M.-H.;
RA "Antibacterial activity of secretolytin, a chromogranin B-derived
RT peptide (614-626), is correlated with peptide structure.";
RL FEBS Lett. 379:273-278 (1996).
CC -!- FUNCTION: Secretogranin I is a neuroendocrine secretory granule
CC protein, which may be the precursor for other biologically active
CC peptides. The 16 pairs of basic AA distributed throughout its
CC sequence may be used as proteolytic cleavage sites.
CC -!- FUNCTION: Secretolytin has antibacterial activity.
CC -!- SUBCELLULAR LOCATION: Neuroendocrine and endocrine secretory
CC granules.
CC -!- PTM: O-glycosylated (Probable).
CC -!- SIMILARITY: Belongs to the chromogranin / secretogranin protein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X55027; CAA38846.1; -;
DR EMBL; U88551; AAC48720.1; -;
DR EMBL; X54489; CAA39109.1; -;
DR PIR; S15901; S15901.
DR InterPro; IPR001819; Chromogranin_AB.
DR InterPro; IPR001990; Granin.
DR Pfam; PF01271; Granin; 1.
DR PRINTS; PR00659; CHROMOGRANIN.
DR PROSITE; PS00422; GRANINS_1; 1.
DR PROSITE; PS00423; GRANINS_2; 1.
DR Sulfation; Cleavage on pair of basic residues; Signal.
KW Sulfation; Cleavage on pair of basic residues; Signal.
FT SIGNAL 1 20
FT CHAIN 21 646
FT PEPTIDE 418 484
FT PEPTIDE 634 646
FT DISULFID 36 57
FT MOD RES 158 158
FT MOD RES 315 315
FT CONFLICT 64 70
FT CONFLICT 70 70
FT CONFLICT 93 98
FT CONFLICT 181 181
FT CONFLICT 261 261
FT CONFLICT 386 386
FT CONFLICT 481 481
FT CONFLICT 597 597
SQ SEQUENCE 646 AA; 73339 MW; 420DB1178FD9E415 CRC64;
Alignment Scores:
Pred. No.: 32.6 Length: 646
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 1 Gaps: 0
US-09-270-437D-7 (1-1946) x SGL_BOVIN (1-646)
QY 1027 CCACGACGCTTACTGGGGCTGCTC 1050
DB 3 PROALALALeuLeuGlyLeuLeu 10
RESULT 14
BACI_MOUSE STANDARD; PRT; 739 AA.
ID BACI_MOUSE
AC P97302;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription regulator protein BACH1 (BTF and CNC homolog 1).
GN BACH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=97042438; PubMed=8887638;
RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA Yamamoto M., Igarashi K.;
RT "Bach proteins belong to a novel family of BTF-basic leucine zipper
RT transcription factors that interact with MafK and regulate
RT transcription through the NF-E2 site.";
RL Mol. Cell. Biol. 16:6083-6095 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Mammary gland;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McWeeny F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Transcriptional regulator that acts as repressor or
CC activator. Binds, in-vitro, to NF-E2 binding sites. play important
CC roles in coordinating transcription activation and repression by
CC MAFK.
CC -!- SUBUNIT: Heterodimer of BACH1 and MAFK.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the bZIP family. CNC subfamily.
CC -!- SIMILARITY: Contains 1 BTF/POZ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC

or send an email to license@isb-sib.ch).

CC EMBL; D86603; BAA13137.1; -.
CC EMBL; BC057894; AAH57894.1; -.
DR HSSP; P34707; 1SKN.
DR TRANSFAC; T04793; -.
DR MGI; MGI:894680; BACH1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005535; F:protein binding; IPI.
DR GO; GO:0003700; P:transcription factor activity; IDA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
DR InterPro; IPR00210; BTF_POZ.
DR InterPro; IPR008917; Bux_transcr_DNA.
DR InterPro; IPR02112; LeuZip_Jun.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00651; BTF; 1.
DR Pfam; PF00170; BZIP; 1.
DR PRINTS; PR00043; LEUZIPPRJUN.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00225; BTF; 1.
DR PROSITE; PS00097; BZIP; 1.
DR PROSITE; PS0217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; Repressor; DNA-binding;
KW Nuclear protein.
FT DOMAIN 34 100 BTF.
FT DNA_BIND 565 580 BASIC MOTIF.
FT DOMAIN 588 610 LEUCINE-ZIPPER.
SQ SEQUENCE 739 AA; 81373 MW; CE2DE606B05F6E32 CRC64;

Alignment Scores:
Pred. No.: 32.1 Length: 739
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.26% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x BAC1_MOUSE (1-739)

OY 1225 ATACGAAGCTTTGGAGTCAGGTGTT 1202
|||||
Db 239 IleArgThrLeuGluSerGlyVal 246

RESULT 15
GYRB_MYXXA
ID GYRB_MYXXA STANDARD; PRT; 815 AA.
AC Q33367;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA gyrase subunit B (EC 5.99.1.13).
GN GYRB.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=ER-15;
RC MEDLINE=98304088; PubMed=9639935;
RX Paitan Y., Boulton N., Ron E.Z., Rosenberg E., Orr E.;
RT "Molecular analysis of the DNA gyrase gene from Myxococcus xanthus.";
RL Microbiology 144:1641-1647(1998).
CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-
stranded DNA in an ATP-dependent manner and also catalyzes the
interconversion of other topological isomers of double-stranded
DNA rings, including catenanes and knotted rings.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
of double-stranded DNA.
CC -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
enzyme forms an A2B2 tetramer.

CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AJ000543; CAA04176.1; -.
DR HSSP; P06982; 1AJ6.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR002288; DNA_gyraseB_C.
DR InterPro; IPR000565; DNA_gyrB.
DR InterPro; IPR001241; DNA_topoisomII.
DR InterPro; IPR006171; Toprim_dom.
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00986; DNA_gyraseB_C; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF01751; Toprim; 1.
DR PRINTS; PR00418; TP12FAMILY.
DR ProDom; PDI49633; DNA_gyraseB_C; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00433; TOP2c; 1.
DR TIGRFAMs; TIGR01059; gyrb; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Topoisomerase; Isomerase; ATP-binding.
SQ SEQUENCE 815 AA; 89636 MW; 3862855FBB805B32 CRC64;

Alignment Scores:
Pred. No.: 31.8 Length: 815
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-7 (1-1946) x GYRB_MYXXA (1-815)

OY 1891 GAGGTCGATCACACCTCAGTGGG 1914
|||||
Db 289 GluGlySerHisLeuSerGly 296

Search completed: August 6, 2004, 13:22:10
Job time : 25.1191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 92.8527 Seconds

(without alignments)
13225.208 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 639
Sequence: 1 gctgtagcggagggtggg.....atttccttcagggttttaaaa 1946

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool_p/US09270437/runat_06082004.141220.1815/app_query.fasta_1.5582
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=oligo806.rspt -MINMATCH=0.1 -DOFCL=0
-LOOPT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09270437.@CGN 1.1 287 /runat_06082004.141220.1815 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL 25: *
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	44.0	577	11	Q8CGX0 rattus norv

2	212	33.2	577	4	Q9NZI8	Q9nzi8 homo sapien
3	166	26.0	577	11	Q8BRH1	Q8brh1 mus musculu
4	166	26.0	577	11	Q84777	Q84777 mus musculu
5	115	18.0	577	11	Q80US9	Q80us9 mus musculu
6	77	12.1	576	13	Q42254	Q42254 gallus gall
7	41	6.4	594	13	Q73932	Q73932 xenopus lae
8	38	5.9	593	13	Q57526	Q57526 xenopus lae
9	36	5.6	579	4	Q00425	Q00425 homo sapien
10	30	4.7	579	11	Q8C2J9	Q8c2j9 mus musculu
11	30	4.7	579	11	Q9CPN8	Q9cpn8 mus musculu
12	24	3.8	582	13	Q9PW80	Q9pw80 brachydanio
13	19	3.0	556	4	Q9Y6M1	Q9y6m1 homo sapien
14	16	2.5	169	11	Q7TF50	Q7tf50 rattus norv
15	15	2.3	545	11	Q7TQF9	Q7tqf9 mus musculu
16	10	1.6	1843	2	Q83WE9	Q83we9 micromonosp
17	9	1.4	322	13	Q9PTU0	Q9ptu0 brachydanio
18	9	1.4	351	16	Q7UXR6	Q7uxr6 rhodopirell
19	9	1.4	423	4	Q9UIN6	Q9uin6 homo sapien
20	9	1.4	424	4	Q9UIN7	Q9uin7 homo sapien
21	9	1.4	424	4	Q9UIP0	Q9uiP0 homo sapien
22	9	1.4	424	4	Q9UBH8	Q9ubH8 homo sapien
23	9	1.4	434	3	Q872V2	Q872v2 neuropsora
24	9	1.4	499	4	Q9NRH5	Q9nrh5 homo sapien
25	9	1.4	499	4	Q9NBH4	Q9nrh4 homo sapien
26	9	1.4	500	4	Q9NPG5	Q9npG5 homo sapien
27	9	1.4	500	4	Q9NRH7	Q9nrh7 homo sapien
28	9	1.4	500	4	Q9NPE0	Q9npe0 homo sapien
29	9	1.4	500	4	Q9NPE2	Q9npE2 homo sapien
30	9	1.4	500	4	Q9NRH8	Q9nrh8 homo sapien
31	9	1.4	500	4	Q9NRH6	Q9nrh6 homo sapien
32	9	1.4	513	4	Q86204	Q86204 homo sapien
33	9	1.4	513	7	Q95IF9	Q95if9 homo sapien
34	9	1.4	528	4	Q8N5P3	Q8n5P3 homo sapien
35	9	1.4	528	4	Q8IZU6	Q8izu6 homo sapien
36	9	1.4	528	4	Q8IZU5	Q8izu5 homo sapien
37	9	1.4	529	4	Q43509	Q43509 homo sapien
38	9	1.4	529	4	Q8IZU4	Q8izu4 homo sapien
39	9	1.4	529	6	Q7YR44	Q7yr44 pan troglod
40	9	1.4	651	13	Q98SM6	Q98sm6 gallus gall
41	8	1.3	55	6	Q46645	Q46645 meles meles
42	8	1.3	77	10	Q8LIP0	Q8lip0 oryza sativ
43	8	1.3	95	16	Q97RK6	Q97rk6 streptococc
44	8	1.3	95	16	Q8CYV6	Q8cyv6 streptococc
45	8	1.3	133	6	Q9GMU9	Q9gmu9 macaca fasc

ALIGNMENTS

RESULT 1	PRELIMINARY	PRT	577 AA
Q8CGX0	Q8CGX0		
ID	Q8CGX0		
AC	Q8CGX0		
DT	01-MAR-2003 (TRENBLrel. 23, Created)		
DT	01-MAR-2003 (TRENBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)		
DE	B-actin zipcode binding protein 1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley;		
RA	Eom T., Singer R.H., Bassell G.J.;		
RT	"Molecular interactions between r2Bp1 and b-actin zipcode required for		
RT	transport of mRNA and stimulation of spine growth.";		
RL	Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AF541940; AAO16210.1;		
DR	GO; GO:0003676; F:nucleic acid binding; IEA.		
DR	InterPro; IPR004087; KH dom.		
DR	InterPro; IPR004088; KH_type 1.		
DR	InterPro; IPR000504; RNA_rec_mot.		
DR	Pfam; PF00013; KH; 4.		

DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50084; KH TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1EE CRC64;

Alignment Scores:

Pred. No.: 1.08e-283 Length: 577
 Score: 281.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 43.97% Indels: 0
 DB: 11 Gaps: 0

US-09-270-437D-7 (1-1946) x Q8CGX0 (1-577)

QY 108 AGGAGCCGGAATAATCAATCCGNAATATTCACCCAGCTCCGATGGGAGTACTGGAC 167
 DB |||||
 QY 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98
 DB |||||
 QY 168 AGCCTGCTGCTCAGTATGCTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
 DB |||||
 QY 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
 DB |||||
 QY 228 ACGCCAGTGTGATGTCACCTATTTCACACCGGAGCAGACCGCAAGCCATCATGAAG 287
 DB |||||
 QY 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
 DB |||||
 QY 288 CTGAATGCCACACAGTGTGAGAGACCATGCCCTCAAGTCTCTACATCCCGATGAGCAG 347
 DB |||||
 QY 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
 DB |||||
 QY 348 ATAGCAGGACCTGAGATGAGCCGCGCGGGGCTTTGGCTCTCGGGTCTCAGCCCGCC 407
 DB |||||
 QY 159 IleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
 DB |||||
 QY 408 CAGGGCTCACCTGCGCAGCGGGGCGCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGC 467
 DB |||||
 QY 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
 DB |||||
 QY 468 CGGCTCTGCTGTCGCCACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
 DB |||||
 QY 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleGlyLysGlyAlaThrIle 218
 DB |||||
 QY 528 CGCAACATCAAAACAGACCCAGTCCAGATAGAGTGCATAGAGAGAGAACCGAGGT 587
 DB |||||
 QY 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
 DB |||||
 QY 588 GCAGCTGMAAAGCCATCAGTGTGCACTCCACCCCTGAGGCTGCTCTCCCTGCTGTAAG 647
 DB |||||
 QY 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
 DB |||||
 QY 648 ATGATCTTGAGATTATGCAAGAGGTAAAGACCAACCAACCGCTGACAGGTCTCC 707
 DB |||||
 QY 259 MetIleLeuGluIleMethHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 DB |||||
 QY 708 CTGAAGATCTGCGCCCATTAATCTTTAGGCGCTCTCATTTGGCAAGGAGGAGCGAAC 767
 DB |||||
 QY 279 LeuLysIleLeuAlaHisAsnAsnPhenValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 DB |||||
 QY 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAATACCATCTCTCTGCTTGCAGACCTT 827
 DB |||||
 QY 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
 DB |||||
 QY 828 ACCCTTTCAACCTCAGAGGACCATCAGTGTGAAGGGGCGCCATCAGAAATTTGTCAGG 887
 DB |||||
 QY 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
 DB |||||
 QY 888 GCCGAGCAGGAATATGAGAAAGTTCGGGAGGCGCTATGAGATGATGCTGCCCTCCCATG 947
 DB |||||
 QY 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
 DB |||||

QY 948 AGC 950
 DB |||||
 DB 359 Ser 359

RESULT 2

Q9NZ18

ID Q9NZ18 PRELIMINARY; PRT; 577 AA.

AC Q9NZ18;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE mRNA-binding protein CRDBP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.

RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
 RA Pantofoakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,
 RA Kittas C., Agnantis N., Pandis N.;
 RT "Ectopic expression of a KH-domain containing protein, highly
 RT homologous to both human IMP-1 and mouse CRD-BP, in benign and
 RT malignant mesenchymal tumors.";
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF198254; AAF37203.1; -.
 DR HSSP; P11940; 1CVJ.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50084; KH TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 577 AA; 63456 MW; 0749A060F25D81D CRC64;

Alignment Scores:

Pred. No.: 1.24e-211 Length: 577
 Score: 212.00 Matches: 212
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.18% Indels: 0
 DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x Q9NZ18 (1-577)

QY 963 CQTGGCCTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCG 1022
 DB |||||
 QY 366 ProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAlaSerSerAlaValProPro 385
 DB |||||
 QY 1023 CTTCCACAGCAGCTTACTGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAG 1082
 DB |||||
 QY 386 ProProSerSerValThrGlyAlaAlaProTyrSerSerPheMetGlnAlaProGluGln 405
 DB |||||
 QY 1083 GAGATGTGTGAGGTGTTTATCCCGCCAGCGAGTGGGCGCCATCATCGCAAGAAGGGG 1142
 DB |||||
 QY 406 GluMetValGlnValPheIleProAlaGlnAlaValGlyAlaIleIleGlyLysLysGly 425
 DB |||||
 QY 1143 CAGCACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACCACCCGAA 1202
 DB |||||
 QY 426 GlnHisIleLysGlnLeuSerArgPheAlaSerAlaSerIleLysIleAlaProGlu 445
 DB |||||
 QY 1203 ACACCTGACTCCAAAGTTGCTATGTTATCATCATCGACCGCCAGAGGCCCAATTCAG 1262
 DB |||||
 QY 446 ThrProAspSerLysValArgMetValIleIleThrGlyProGluAlaGlnPheLys 465
 DB |||||
 QY 1263 GCTCAGGAGAGATCTATGGCAACTCAAGGAGGAGAACTTCTTGTCCCAAGGAGGAA 1322
 DB |||||
 QY 466 AlaGlnGlyArgIleThrGlyLysLeuLysGluGluAsnPhenGlyProLysGluGlu 485
 DB |||||

```
QY 1323 GTGAAGCTGGACACCCACATACGTTGTCCAGCATCAGCAGCTGGCCGGTCAATTGGCAAA 1382
|
|
|
Db 486 VallysLeuGluThrHisileargValProAlaSerAlaAlaGlyArgValileGlyLys 505
|
|
|
QY 1383 GGTGGAAAACCGTGAACGAGTTGCAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAGA 1442
|
|
|
Db 506 GlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluValValProArg 525
|
|
|
QY 1443 GACCAGACCCCTGATGAGAACACACAGGTCTATCGTGAATAATCATCGGCATTTTATGCC 1502
|
|
|
Db 526 AspGlnThrProAspGluAsnAspGlnValileVallylelleGlyHisPheTyra 545
|
|
|
QY 1503 AGTCAGATCGCTCAACCGGAAGATCCGAGACATCCTGGCCCGAGTTAAGCAGCAGCATCAG 1562
|
|
|
Db 546 SerGlnMetAlaGlnArgLysileargAspIleLeuAlaGlnVallylelleGlnHisGln 565
|
|
|
QY 1563 AAGGACACAGTAACACAGCCCGCAGGACGAGGAGGAAG 1598
|
|
|
Db 566 LysGlyGlnSerAsnGlnAlaGlnAlaArgA-Glys 577
```

RESULT 3

```
Q8BRH1
ID Q8BRH1 PRELIMINARY; PRT; 577 AA.
AC C8BRH1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGF2BP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK044850; BAC32119.1; -.
DR MGD; MGI:1890357; IGF2bp1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D75897450841532E CRC64;
```

Alignment Scores:

```
Pred. No.: 1.35e-163 Length: 577
Score: 166.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.98% Indels: 0
DB: 11 Gaps: 0
```

US-09-270-437D-7 (1-1946) x Q8BRH1 (1-577)

```
QY 453 GTGAGATCCCTCGGCTCGTCCGCCACCCAGTGTGGTCCCATTTGGCAAG 512
|
|
|
Db 194 ValAspIleProLeuArgLeuLeuValProThrGlnTyValGlyAlaIlelleGlyLys 213
|
|
|
QY 513 GAGGGGCCACCATCGGCCACATCACAAAACACACCCAGTCCAAAGATAGACGTGCATAGG 572
|
|
|
```

```
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
|
|
|
QY 573 AAGAGAACCGAGTGGAGCTGAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGCTGC 632
|
|
|
Db 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisThrProGluGlyCys 253
|
|
|
QY 633 TCCCTCCCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACG 692
|
|
|
Db 254 SerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThr 273
|
|
|
QY 693 GCTGACGAGTTCCCTGGAAGATCCCTGGCCCATTAATACTTTGTAGGGCTCTCATTTGGC 752
|
|
|
Db 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
|
|
|
QY 753 AAGCAAGACGCGACCTGGAAGAGGTAGACACAGATACCGAGACAAAATCACCATCTCC 812
|
|
|
Db 294 LysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIleSer 313
|
|
|
QY 813 TC GTTGAAGACCTTACCCCTTTACAACCTGAGAGGACCATCATCTGTGAAGGGGGCCATC 872
|
|
|
Db 314 SerLeuGlnAspLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAlaIle 333
|
|
|
QY 873 GAGAATTGTTCCAGGGCCGAGCAGGAATAATGAAGAAAGTTCCGGAGGCTATGAGAAT 932
|
|
|
Db 334 GluAsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyArgLys 353
|
|
|
QY 933 GATGTGGCTGCCATGAGC 950
|
|
|
Db 354 AspValAlaAlaMetSer 359
|
|
|
RESULT 4
Q88477
ID Q88477 PRELIMINARY; PRT; 577 AA.
AC Q88477;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coding region determinant binding protein.
GN IGF2BP1 OR CRDBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92217743; PubMed=1559612;
RA Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
RT "Control of c-myc mRNA half-life in vitro by a protein capable of
binding to a coding region stability determinant.";
RL Genes Dev. 6:642-654(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94158886; PubMed=8114742;
RA Herrick D.J., Ross J.;
RT "The half-life of c-myc mRNA in growing and serum-stimulated cells:
influence of the coding and 3' untranslated regions and role of
ribosome translocation.";
RL Mol. Cell. Biol. 14:2119-2128(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94179348; PubMed=8132663;
RA Prokipcak R.D., Herrick D.J., Ross J.;
RT "Purification and properties of a protein that binds to the C-terminal
coding region of human c-myc mRNA.";
RL J. Biol. Chem. 269:9261-9269(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97322234; PubMed=9178888;
RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
RA Gruppiso P.A., Ross J.;
RT "Developmental regulation of CRD-BP, an RNA-binding protein that
stabilizes c-myc mRNA in vitro.";
RL Oncogene 14:1279-1286(1997).
```

```

[5]
RP SEQUENCE FROM N.A.
RA Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
RA Fleisig A.J.,
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF061569; AAC72743.1; -.
DR EMBL; AK013940; BAB29071.1; -.
DR HSSP; P11940; ICVJ.
DR MGD; MGI:1890357; Igfbp1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63450 MW; EFBBLAP2FF9F0344 CRC64;

Alignment Scores:
Pred. No.: 1.35e-163 Length: 577
Score: 166.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.98% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-7 (1-1946) x 088477 (1-577)

QY 453 GTGACATCCCTCGCTCGCTCGTCCACCCAGCATGTGGTCCCATTTGSCAAG 512
DB 194 ValAspIleProLeuArgLeuLeuValProThrGlnTyValGlyAlaIleLeuGlyLys 213
QY 513 GAGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGATAGG 572
DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACCCAGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTCAGGGCTGC 632
DB 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHiserThrProGluGlyCys 253
QY 633 TCTCCGCTGTGAATGATCTTTGGAGATTATGCATAAAGAGCTTAAGACACCAAAACG 692
DB 254 SerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThr 273
QY 693 GCTGACGAGGTTCCCTGAGATCTCTGGCCCATATAAATCTTGTAGGGGCTCTCATTCGC 752

```

```

DB 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACCGACCTGAAGAGGTAGAGCAAGTACCGAGACCAAAAATCACCATCTCC 812
DB 294 LysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTTACCTTTTCAACCCCTGAGAGGACCATCACTGTGAGGGGGCCATC 872
DB 314 SerLeuGlnAspLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAlaIle 333
QY 873 GAGAAATTTCGAGGCGCGACGACGAGAAATAATAGAAAGTTCGGGAGGCTTATGAGAAAT 932
DB 334 GluAsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyArgGluAsn 353
QY 933 GATGTGGCTGCATCAGC 950
DB 354 AspValAlaAlaMetSer 359

RESULT 5
Q80US9 PRELIMINARY; PRT; 577 AA.
AC Q80US9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Igfbp1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.,
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.,
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.,
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heish F.,
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.,
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.,
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.,
RA Raha S.S.; Lequellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.,
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.,
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.,
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.,
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.,
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.,
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.,
RA Krzywinski M.I.; Skalska U.; Smallos D.E.; Schnerch A.; Schein J.B.,
RA Jones S.J.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051679; AAH51679.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D439F544257DA3CE CRC64;

```

Alignment Scores:		2.47e-110	Length: 577
Pred. No.:	115.00	Matches: 115	
Score:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	18.00%	Indels: 0	
DB:	11	Gaps: 0	
US-09-270-437D-7 (1-1946) x Q80US9 (1-577)			
QY	1233	ATCACTGGACCGCCAGAGCCCAATTCAGGCTCAGGAGAGATCTATGGCAACTCAAG	1292
DB	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeuLys	475
QY	1293	GAGGAGAACTCTCTTGGTCCCAAGAGAGAGTGAAGCTGGACACCCACATACGTGTGCCA	1352
DB	476	GluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValPro	495
QY	1353	GCATCAGCAGCTGGCCGGTCAATTCGCAAGAGTGGAAACCGTGAACGAGTTCAGAAAT	1412
DB	496	AlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
QY	1413	TTGACGGCAGCTGAGTGGTGTAGTACCAAGAGACGACCCCTGATCAGAACACGAGTTC	1472
DB	516	LeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnVal	535
QY	1473	ATCGTGAATAATCATCGGACATTTCTATGCCAGTCAAGTCCAGGCTCAACGGGAAGATCCGAGAC	1532
DB	536	IleValLysIleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLysIleArgAsp	555
QY	1533	ATCTGCCCCAGTTTAAGCAGCAGCATCAGAGGGCAGAGTAAC	1577
DB	556	IleLeuAlaGlnValLysGlnGlnHisGlnLysGlyGlnSerAsn	570
RESULT 6			
ID	042254	PRELIMINARY;	PRT; 576 AA.
AC	042254;		
DT	01-JAN-1998 (T-EMBLrel. 05, Created)		
DT	01-AUG-1998 (T-EMBLrel. 05, Last sequence update)		
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)		
DE	Zipcode-binding protein.		
GN	ZBP1.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9720007; PubMed=9121465;		
RA	Ross A.F., Oleynikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;		
RT	"Characterization of a beta-actin mRNA zipcode-binding protein."		
RL	Mol. Cell. Biol. 17:2158-2165(1997).		
CC	-1- SIMILARITY: CONTAINS 4 KH DOMAINS.		
DR	EMBL; AF026527; AAB82295.1; -		
DR	GO; GO:0003676; F:nucleic acid binding; IEA.		
DR	InterPro; IPR004087; KH_dom.		
DR	InterPro; IPR000504; RNA_rec_mot.		
DR	Pfam; PF00013; KH; 4.		
DR	Pfam; PF00076; xrm; 2.		
DR	SMART; SM00322; KH; 4.		
DR	SMART; SM00360; RRM; 2.		
DR	PROSITE; PS00084; KH_TYPE_1; 4.		
DR	PROSITE; PS0102; RRM; 2.		
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.		
SQ	SEQUENCE 576 AA; 63271 MW; 01AAF2D1D81C8811 CRC64;		
Alignment Scores:			
Pred. No.:	1.2e-70	Length: 576	
Score:	77.00	Matches: 77	
Percent Similarity:	100.00%	Conservative: 0	
US-09-270-437D-7 (1-1946) x O73932 (1-594)			
QY	1221	CGTATGTTTATCATCTACCTGACCGCCAGAGGCCCAATTCAGGCTCAGGAGAAATCTAT	1280

Best Local Similarity:		100.00%	Mismatches: 0
Query Match:	12.05%	Indels: 0	
DB:	13	Gaps: 0	
US-09-270-437D-7 (1-1946) x O42254 (1-576)			
QY	1233	ATCACTGGACCGCCAGAGCCCAATTCAGGCTCAGGAGAGATCTATGGCAACTCAAG	1292
DB	455	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeuLys	474
QY	1293	GAGGAGAACTCTCTTGGTCCCAAGAGAGAGTGAAGCTGGACACCCACATACGTGTGCCA	1352
DB	475	GluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValPro	494
QY	1353	GCATCAGCAGCTGGCCGGTCAATTCGCAAGAGTGGAAACCGTGAACGAGTTCAGAAAT	1412
DB	495	AlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	514
QY	1413	TTGACGGCAGCTGAGTGGTGTAGTACCAAGAGACGACCCCTGATCAGAAAC	1463
DB	515	LeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsn	531
RESULT 7			
ID	073932	PRELIMINARY;	PRT; 594 AA.
AC	073932;		
DT	01-AUG-1998 (T-EMBLrel. 07, Created)		
DT	01-AUG-1998 (T-EMBLrel. 07, Last sequence update)		
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)		
DE	VGI RNA binding protein variant D.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Havin L., Git A., Elisha Z., Oberman F., Yaniv K.;		
RA	Pressman Schwartz S., Standart N.M., Yisraeli J.K.;		
RL	Genes Dev. 0:0-0(1998).		
[2]			
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98228351; PubMed=9560341;		
RA	Deshler J.O., Hightett M.I., Abramson T., Schnapp B.J.;		
RT	"A highly conserved RNA-binding protein for cytoplasmic mRNA		
RT	localization in vertebrates."		
RL	Curr. Biol. 8:489-496(1998).		
CC	-1- SIMILARITY: CONTAINS 4 KH DOMAINS.		
DR	EMBL; AF064634; AAC18598.1; -		
DR	EMBL; AF055923; AAC41285.1; -		
DR	GO; GO:0003676; F:nucleic acid binding; IEA.		
DR	InterPro; IPR004087; KH_dom.		
DR	InterPro; IPR004088; KH_type_1.		
DR	InterPro; IPR000504; RNA_rec_mot.		
DR	Pfam; PF00013; KH; 4.		
DR	Pfam; PF00076; xrm; 2.		
DR	SMART; SM00322; KH; 4.		
DR	SMART; SM00360; RRM; 2.		
DR	PROSITE; PS00084; KH_TYPE_1; 4.		
DR	PROSITE; PS0102; RRM; 2.		
SQ	SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;		
Alignment Scores:			
Pred. No.:	4.7e-33	Length: 594	
Score:	41.00	Matches: 41	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	6.42%	Indels: 0	
DB:	13	Gaps: 0	

```

Db 465 ArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyr 484
QY 1281 GGCACAACTCAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCAC 1340
D 485 GlyLysLeuLysGluGluAsnPheGlyProLysGluGluValLysLeuGluThrHis 504
QY 1341 ATA 1343
D 505 Ile 505

RESULT 8
O57526 PRELIMINARY; PRT; 593 AA.
AC O57526;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KH domain-containing transcription factor B3.
OS Xenopus laevis (African clawed frog)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92249652; PubMed=1577195;
RA Pfaff S.L., Taylor W.L.;
RT "Characterization of a Xenopus oocyte factor that binds to a
RT developmentally regulated cis-element in the TFIIF gene.";
RL Dev. Biol. 151:306-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Griffin D., Taylor W.L.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL Genes Dev. 0:0-0(1998).
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF042353; AAB97457.1; -.
DR EMBL; AF064633; AAC18597.1; -.
DR GO; GO:0003676; F: nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type 1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; zrm; 2.
DR SMART; SMO0322; KH; 4.
DR SMART; SMO0360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP 1; FALSE NEG.
SQ SEQUENCE 593 AA; 65385 MW; 5A5E4B4A1D55DF7 CRC64;

Alignment Scores:
Pred. No.: 6,38e-30 Length: 593
Score: 38.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.95% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-7 (1-1946) x O57526 (1-593)

QY 1221 CGTATGTTATCATCTACTGACCGCAGAGGCCCAATTCAGCTCAGGAAGAATCTAT 1280
D 464 ArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyr 483
QY 1281 GGCACAACTCAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAG 1334
D 484 GlyLysLeuLysGluGluAsnPheGlyProLysGluGluValLysLeuGlu 501

```

```

RESULT 9
O00425 PRELIMINARY; PRT; 579 AA.
AC O00425;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative RNA binding protein KOC (KOC).
GN KOC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;
RL Oncogene 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Mueller-Pillasch F., Lacher U., Wallrapp C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; U97188; AAC35208.1; -.
DR EMBL; U76705; AAD09223.1; -.
DR GO; GO:0005737; C: cytoplasm; TAS.
DR GO; GO:0003723; F: RNA binding; TAS.
DR GO; GO:0007345; P: embryogenesis and morphogenesis; TAS.
DR GO; GO:0006412; P: protein biosynthesis; TAS.
DR GO; GO:0006396; P: RNA processing; TAS.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; zrm; 2.
DR SMART; SMO0322; KH; 4.
DR SMART; SMO0360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP 1; FALSE NEG.
SQ SEQUENCE 579 AA; 63720 MW; AB5C3A8EE3C135C5 CRC64;

Alignment Scores:
Pred. No.: 7,85e-28 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-7 (1-1946) x O00425 (1-579)

QY 492 GTGGGTGCATTATTGGCAAGGAGGGCCACCATCGCAACATCCAAACAGACCCAG 551
D 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 552 TCCAAGATAGCGTCATAGGAAGGAGACGCGAGTGCAGCTCAAAA 599
D 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 10
Q8C2J9 PRELIMINARY; PRT; 579 AA.
ID Q8C2J9
AC Q8C2J9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGF2BF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD; TISSUE=Thymus;
 RA MEDLINE=22354683; PubMed=12466851;
 RX The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT 'Analysis of the mouse transcriptome based on functional annotation of
 RL 60,770 full-length cDNAs.';
 RL Nature 420:563-573(2002).
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63551 MW; 937B501A95D06B77 CRC64;

 Alignment Scores:
 Pred. No.: 1.45e-21 Length: 579
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 11 Gaps: 0

 US-09-270-437D-7 (1-1946) x Q8C2J9 (1-579)
 Qy 492 GTGGGTGCCATTATTGCAAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValglyalalleleGlyLysGluGlyAlaThrIleargasullehrlysglnthrln 226
 Qy 552 TCCAGATACAGTGCATAGAGGAGAAC 581
 Db 227 SerLysileAspValHisArgLysGluasn 236

 RESULT 11
 Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library.
 DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGf2BP3 OR 2610101N1LRK OR NIMP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217811;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scaulli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT 'Functional annotation of a full-length mouse cDNA collection.';
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bontalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnarch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT 'Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.';
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -.
 DR EMBL; AB046173; BAB19755.1; -.
 DR EMBL; BC045138; AAH45138.1; -.
 DR EMBL; BC049082; AAH49082.1; -.
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

 Alignment Scores:
 Pred. No.: 1.45e-21 Length: 579
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 11 Gaps: 0

 US-09-270-437D-7 (1-1946) x Q8CPN8 (1-579)
 Qy 492 GTGGGTGCCATTATTGCAAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValglyalalleleGlyLysGluGlyAlaThrIleargasullehrlysglnthrln 226
 Qy 552 TCCAGATACAGTGCATAGAGGAGAAC 581
 Db 227 SerLysileAspValHisArgLysGluasn 236

 RESULT 11
 Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library.
 DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGf2BP3 OR 2610101N1LRK OR NIMP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217811;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scaulli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT 'Functional annotation of a full-length mouse cDNA collection.';
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bontalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnarch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT 'Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.';
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -.
 DR EMBL; AB046173; BAB19755.1; -.
 DR EMBL; BC045138; AAH45138.1; -.
 DR EMBL; BC049082; AAH49082.1; -.
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

 Alignment Scores:
 Pred. No.: 1.45e-21 Length: 579
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 11 Gaps: 0

 US-09-270-437D-7 (1-1946) x Q8CPN8 (1-579)
 Qy 492 GTGGGTGCCATTATTGCAAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValglyalalleleGlyLysGluGlyAlaThrIleargasullehrlysglnthrln 226
 Qy 552 TCCAGATACAGTGCATAGAGGAGAAC 581
 Db 227 SerLysileAspValHisArgLysGluasn 236

 RESULT 11
 Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library.
 DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGf2BP3 OR 2610101N1LRK OR NIMP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217811;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scaulli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT 'Functional annotation of a full-length mouse cDNA collection.';
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bontalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnarch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT 'Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.';
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -.
 DR EMBL; AB046173; BAB19755.1; -.
 DR EMBL; BC045138; AAH45138.1; -.
 DR EMBL; BC049082; AAH49082.1; -.
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

 Alignment Scores:
 Pred. No.: 1.45e-21 Length: 579
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 11 Gaps: 0

 US-09-270-437D-7 (1-1946) x Q8CPN8 (1-579)
 Qy 492 GTGGGTGCCATTATTGCAAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValglyalalleleGlyLysGluGlyAlaThrIleargasullehrlysglnthrln 226
 Qy 552 TCCAGATACAGTGCATAGAGGAGAAC 581
 Db 227 SerLysileAspValHisArgLysGluasn 236

 RESULT 11
 Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library.
 DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGf2BP3 OR 2610101N1LRK OR NIMP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217811;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scaulli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT 'Functional annotation of a full-length mouse cDNA collection.';
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bontalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnarch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT 'Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.';
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -.
 DR EMBL; AB046173; BAB19755.1; -.
 DR EMBL; BC045138; AAH45138.1; -.
 DR EMBL; BC049082; AAH49082.1; -.
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

 Alignment Scores:
 Pred. No.: 1.45e-21 Length: 579
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 11 Gaps: 0

 US-09-270-437D-7 (1-1946) x Q8CPN8 (1-579)
 Qy 492 GTGGGTGCCATTATTGCAAGGAGGGGCCACCATCGCAACATCACAAACAGACCCAG 551
 Db 207 ValglyalalleleGlyLysGluGlyAlaThrIleargasullehrlysglnthrln 226
 Qy 552 TCCAGATACAGTGCATAGAGGAGAAC 581
 Db 227 SerLysileAspValHisArgLysGluasn 236

 RESULT 11
 Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library.
 DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGf2BP3 OR 2610101N1LRK OR NIMP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217811;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scaulli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT 'Functional annotation of a full-length mouse cDNA collection.';
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bontalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnarch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT 'Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.';
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -.
 DR EMBL; AB046173; BAB19755.1; -.
 DR EMBL; BC045138; AAH45138.1; -.
 DR EMBL; BC049082; AAH49082.1; -.
 DR MGD; MGI:1890359; Igf2bp3.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_

AC Q7TP50;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ab2-255.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.P.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY225199; AAP99600.1; -; 19BF6295C1000CEA CRC64;
SQ SEQUENCE 169 AA; 17743 MW; 19BF6295C1000CEA CRC64;

Alignment Scores:
Pred. No.: 7.13e-07 Length: 169
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-7 (1-1946) x Q7TP50 (1-169)

QY 1215 AAGTTCGTATGGTTATCATCTGACGCCAGCCAGCGCCCAATTCAAG 1262
Db 59 LysValArgMetValIleIleThrGlyProGluAlaGlnPheLys 74

RESULT 15
Q7TQF9 PRELIMINARY; PRT; 545 AA.
AC Q7TQF9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Baha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX Strausberg R.;

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC054552; AAH54552.1; -;
KW Hypothetical protein.
SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;

Alignment Scores:
Pred. No.: 6.75e-06 Length: 545
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.35% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-7 (1-1946) x Q7TQF9 (1-545)

QY 732 TTTCAGGCGCTCTCATTCGCAAGGACGCGACCTGAAGAG 776
Db 277 PheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLys 291

Search completed: August 6, 2004, 13:30:44
Job time : 104.853 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 6, 2004, 13:12:26 ; Search time 144.09 Seconds
(without alignments)
12875.302 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaatgtttattt 3283

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues
Word size: 1

Total number of hits satisfying chosen parameters: 2987630
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DEV=xlp
-Q/cgnt2_1/USFTO_spool_p/US0270437/runat_06082004_141220_1795/app.query.fasta_1.5582
-DB=A.Geneseq_29Jan04 -QMT=fastan -SURFIX=oligo806.rag -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0270437 @CGN 1.1.185 @runat_06082004_141220_1795
-NCPUI=6 -ICPU=3 -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOF=6 -FGAPEXT=7 -YGAPOF=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	577	53.6	594	ABG06795	Abg06795 Novel hum
2	556	51.6	556	ABG96346	Abg96346 Human ova
3	361	33.5	614	ABG06794	Abg06794 Novel hum
4	357	33.1	620	AAU16163	AAU16163 Human nov
5	357	33.1	620	ABU55232	ABU55232 Human nov
6	160	14.9	171	AAU16166	AAU16166 Human nov
7	160	14.9	171	ABU55235	ABU55235 Human nov
8	97	9.0	97	AAG03261	Aag03261 Human sec
9	70	6.5	192	AAU16164	AAU16164 Human nov
10	70	6.5	192	ABU55233	ABU55233 Human nov

11	66	6.1	171	4	AAU16583	AAU16583 Human nov
12	66	6.1	171	6	ABU55652	ABU55652 Human nov
13	44	4.1	555	6	ABU9799	ABU9799 Novel hum
14	32	3.0	187	6	AAO23971	AAO23971 Human IGF
15	20	1.9	47	4	AAU16581	AAU16581 Human nov
16	20	1.9	47	6	ABU55650	ABU55650 Human nov
17	19	1.8	148	4	ABG21962	ABG21962 Novel hum
18	19	1.8	250	4	AAU16579	AAU16579 Human nov
19	19	1.8	250	6	ABU55648	ABU55648 Human nov
20	19	1.8	261	4	AAU16161	AAU16161 Human nov
21	19	1.8	261	6	ABU55230	ABU55230 Human nov
22	19	1.8	319	4	AAU93826	AAU93826 Human pol
23	19	1.8	577	2	AAU30649	AAU30649 A murine
24	19	1.8	579	3	AAU11328	AAU11328 Human lun
25	19	1.8	579	3	AAU11365	AAU11365 Human lun
26	19	1.8	579	5	ABU74960	ABU74960 Human lun
27	19	1.8	579	5	ABU75053	ABU75053 Human lun
28	19	1.8	579	5	ABU74997	ABU74997 Human lun
29	19	1.8	579	5	ABU75054	ABU75054 Human lun
30	19	1.8	579	5	ABU61917	ABU61917 Human lun
31	19	1.8	579	5	ABU61974	ABU61974 Human lun
32	19	1.8	579	5	ABU61880	ABU61880 Human lun
33	19	1.8	579	5	ABU61973	ABU61973 Human lun
34	19	1.8	579	7	ADA28536	ADA28536 Recombina
35	19	1.8	579	7	ADA28539	ADA28539 Recombina
36	19	1.8	579	7	ADA28438	ADA28438 Human lun
37	19	1.8	579	7	ADA28266	ADA28266 Human lun
38	19	1.8	579	7	ADA14066	ADA14066 Human src
39	19	1.8	579	7	ADE53471	ADE53471 Human lun
40	19	1.8	586	5	ABU75048	ABU75048 Human lun
41	19	1.8	586	5	ABU61968	ABU61968 Human lun
42	19	1.8	586	7	ADA28517	ADA28517 Recombina
43	18	1.7	619	4	ABG21963	ABG21963 Novel hum
44	17	1.6	583	4	ABG12592	ABG12592 Novel hum
45	15	1.4	20	5	ABU75062	ABU75062 Human lun

ALIGNMENTS

RESULT 1
ABG06795
ID ABG06795 standard; protein; 594 AA.
XX
AC ABG06795;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5786.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS70982.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

PS Claim 20; SEQ ID NO 37154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 594 AA;

Alignment Scores:

Pred. No.:	0	Length:	594
Score:	577.00	Matches:	577
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	53.57%	Indels:	0
DB:	4	Gaps:	0

US-09-270-437D-8 (1-3283) x ABG06795 (1-594)

Qy	7	GGAGGAGCGGAGGAGCGCGGTACCGGGCGGGGAGCGCGGTCTCGGGAGAGA	66
Db	18	GlyGlyGlyGluGluArgValProGlyArgGlySerArgGlyLeuSerGlyLysArg	37
Qy	67	CGGATGATGAACAGCTTACATCGGACCTGAGCCCGCGTCCACCCCGGACGCTC	126
Db	38	ArgMetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeu	57
Qy	127	CGGAGCTCTTTGGGACAGGAAGTGGCCCTCGCGGGACAGTCTCGTGAAGTCCGGC	186
Db	58	ArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGly	77
Qy	187	TAGCCCTCTGTGACTACCCGACAGAACTGGGCCATCGGGCCATCGAGACCTCTCG	246
Db	78	TyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSer	97
Qy	247	GGTAAATGGAATTCATGGGAAATCATCGAAGTTGATTACTCAGTCTCTAAAGCTA	306
Db	98	GlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeu	117
Qy	307	AGGAGCAGGAATTCAGATTCGAACATCCCTCTCTCACTCGAGTGGAGGTGTGGAT	366
Db	118	ArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAsp	137
Qy	367	GGACTTTTGGCTCAATATGGACAGTGGAGAATGTGGAACAAGTCAACACACACAGAA	426
Db	138	GlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGlu	157
Qy	427	ACCGCCCTGTCACTCATATGCAACAGAGAGAGAGAAATACCCATGAGGAGAG	486
Db	158	ThrAlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLys	177
Qy	487	CTAAGCGGCGCATCAGTTGAGACTACTCTCTCAGATTTCCTACATCCCGATGAGAG	546
Db	178	LeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGlu	197

Qy	547	GTGAGCTCCCTTCCGCCCTCAGGAGCCCGCGTGGGACCACTCTTCCCGGAGCAA	606
Db	198	ValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGln	217
Qy	607	GGCCACGCCCTGGGGGACATCTTCAGGCCACAGAGATTGATTTCCGGCTCGCATCTG	666
Db	218	GlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeu	237
Qy	667	GTCCCAACCCAGTTTGTGTGTCATCATCGAAAGAGGGCTTGACCATAAAGAAATC	726
Db	238	ValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIle	257
Qy	727	ACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAG	786
Db	258	ThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGlu	277
Qy	787	AAGCTGTCCATCCATGCCACCCAGAGGGGACTTCTGAAGCATCCCATGATCTT	846
Db	278	LysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeu	297
Qy	847	GAAATCATGCAAGAGGGCAGATGAGCCAACTAGCCGAGAGATTCCTCTGAAATC	906
Db	298	GluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIle	317
Qy	907	TTGGCACAATGGCTTGGTGGAGAGACTGATTGGAAAAGAGCGAGAAATTTGAAGAA	966
Db	318	LeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyAsnLeuLysLys	337
Qy	967	ATTGACATGAAACAGGAGCCAGATTAACATCTCATCTTTCAGGATTTGAGCATATAC	1026
Db	338	IleGluHisGluThrGlyThrLysIleIleSerSerLeuGlnAspLeuSerIleTyr	357
Qy	1027	AACCCGAAAGAACCATCACTGTGAAGGCGACAGTTCAGGCGCTGTGCGAGTGTAGATA	1086
Db	358	AsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIle	377
Qy	1087	GAGATTATGAAGAGCTCGGTGAGGCTTTGAAATGATATGCTGGCTGTATACACCCAC	1146
Db	378	GluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnThrHis	397
Qy	1147	TCCGGATACTTCTCAGCCTGTACCCCATCACCAGTTTCGCCGTTCGCCCATCATCAC	1206
Db	398	SerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHisHis	417
Qy	1207	TCTTATCCAGACGAGATGTGAATCTCTTATCCCAACCCAGGCTGTGGCGGCATC	1266
Db	418	SerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIle	437
Qy	1267	ATCGGAGAGGGGGCACACATCAACAGCTGGCGAGATTCGCCGAGCTCTATCAAG	1326
Db	438	IleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLys	457
Qy	1327	ATTGCCCTTCGGGAAGGCCAGAGCTCAGCGAAGATGTGTCTCATCATCCCGGCCACCG	1386
Db	458	IleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProPro	477
Qy	1387	GAACCCCATTTCAAGGCCCGAGGACGATCTTTGGGAACTTGAAGAGGAAACTCTTT	1446
Db	478	GluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhePhe	497
Qy	1447	AACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGGCTCTTCCACAGCTGGC	1506
Db	498	AsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGly	517
Qy	1507	CGGTGATTTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAACCATGAGAA	1566
Db	518	ArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGlu	537
Qy	1567	GTCTATCGTCTGTGACCAACAGCCAGATGAAATGAGGAAGTGTATCGTCAGAAATATC	1626
Db	538	ValIleValProArgAspGlnThrProAspGluAsnGluValIleValArgIleIle	557

1627 GGGCACTCTTTGCTAGCCAGCACTGCCAGCCAGGATCAGGGAATTTGTACACAGGTG 1686
 558 GllHisPhePheAlaSerGlnThrAlaGlnArgLysileArgLulleValGlnGlnVal 577
 1687 AAGCAGCAGGAGCAGAAATACCTCAGGAGTCCCTCAGCAGCGCAGCAG 1737
 578 LysGlnGlnGlnGlnLysTyProGlnGlnValAlaSerGlnArgSerLys 594

RESULT 2
 ABG96346
 ID ABG96346 standard; protein; 556 AA.
 AC ABG96346;
 XX
 XX
 DT 11-DEC-2002 (first entry)
 DE Human ovarian cancer marker M452.
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nontuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker.
 OS Homo sapiens.
 XX
 XX
 PN WO200271928-A2.
 XX
 XX
 PD 19-SEP-2002.
 XX
 XX
 PF 14-MAR-2002; 2002WO-US007826.
 XX
 XX
 PR 14-MAR-2001; 2001US-0276025P.
 PR 14-MAR-2001; 2001US-0276026P.
 PR 10-AUG-2001; 2001US-0311732P.
 PR 19-SEP-2001; 2001US-0323580P.
 PR 26-SEP-2001; 2001US-0324967P.
 PR 26-SEP-2001; 2001US-0325102P.
 PR 26-SEP-2001; 2001US-0325149P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Monahan JE, Ganavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrisey MF, Olandt PU, Sen A, Vieby PO, Mills GB;
 PI Bast RC, Lu K, Schmandt RE, Zhao X, Giatt K;
 XX
 DR WPI; 2002-723277/78.
 DR N-PSDB; ABS76442.
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient.
 XX
 PS Disclosure; Page 263-264; 48ipp; English.
 XX
 CC The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterising cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),

testicular disorders (e.g. nontuberculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining
 CC whether ovarian cancer has metastasized or is likely to metastasize,
 CC selecting a composition for inhibiting ovarian cancer, assessing the
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
 CC cancer or at risk of developing ovarian cancer. The present amino acid
 CC sequence represents one of the ovarian cancer markers described in the
 CC invention
 XX
 XX Sequence 556 AA;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 556
 Score: 556.00 Matches: 556
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.62% Indels: 0
 DB: 5 Gaps: 0

US-09-270-437D-8 (1-3283) X ABG96346 (1-556)

QY 70 ATGATGAACAAAGCTTTATCATCGGGAACCTGAGCCCGCGCTCACCGCCGACGACCTCCGG 129
 Db 1 MetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
 QY 130 CAGCTCTTTGGGACAGGAAGCTGCCCTCGCGGACAGTCTCTGCTGAAGTCCGGCTAC 189
 Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyLysValLeuLeuLysSerGlyTyr 40
 QY 190 GCCTTCGTGACTACCCCGACAGAACTGGGCCCATCGGCCCATCGAGACCTCTCGGGT 249
 Db 41 AlaPheValAspTyrProAspGlnAsnTTPAlaIleArgAlaIleGluThrLeuSerGly 60
 QY 250 AAGTGAANTGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
 Db 61 LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
 QY 310 AGCAGGAAATTCAGATTCCGAAACATCCCTCTCACTGAGTGGGAGGTGTGGATGGA 369
 Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
 QY 370 CTITTTGCTCAATATGGGACAGTGGAGATGGGAACAAGTCAACACACACACAGAAACC 429
 Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGlnValAsnThrAspThrGluThr 120
 QY 430 GCCCTTCTCACTCATATGCAATATGCAACAGAGAAAGCAAAATAGCCATGAGAGCTA 489
 Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
 QY 490 AGCGGGATCAGTTTGGAAGTACTCTCTCAAGATTCTTACATCCCGGATGAGAGGTG 549
 Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
 QY 550 AGCTCCCTTCGCGCCCTCAGCGAGCCAGCGTGGGAGCACCTCTTCCCGGAGCAAGGC 609
 Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
 QY 610 CAGCCCTCGGGGCGCTTCTCAGGCCAGACAGATTGATTCCCGCTCGGGATCTCTGGTC 669
 Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
 QY 670 CCCACCCAGTTTCTGGTGCATCATCGGAAGAGGGCTTCACCAATAAGAACATCACT 729
 Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
 QY 730 AAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCCAGAGAG 789
 Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
 QY 790 CCGTGTACCATCATCCACCCCGAGAGGGAGCTTCTTGAAGCATGCCGATGATCTTGA 849

Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
 QY 850 ATCATGCGAAGAGGCGAGTACAGCAACCTAGCCGAGAGATCTCTGAAATCTTG 909
 Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluProLeuLysIleLeu 280
 QY 910 GCACCAATGCTTGGTGAAGACTGATTGAAAGAGCAGAAATTTGAAGAAATT 969
 Db 281 AlahisnGlyLeuValGlyArgLeuIleGlyLysGluArgAsnLeuLysIle 300
 QY 970 GAACATGAACAGGCGACAGATTAACATCTCATCTTTGAGGATTTGACATATCAAC 1029
 Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
 QY 1030 CCGGAAAGAACCATCACTGTGAGGCGACAGTTGAGCCCTGCTGCTGAGATAGAG 1089
 Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
 QY 1090 ATTATGAAGAGTGCCTGAGGCGCTTTGAAATCATATGCTGCTGTAAACCCACTCC 1149
 Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnThrHisSer 360
 QY 1150 GGATACCTTCCAGCTGTACCCCATCAGAGTTTGGCCGCTTCCGCTATCATCTCT 1209
 Db 361 GlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHisSer 380
 QY 1210 TATCAGAGCAGAGATGTGAATCTCTTCTCATCCCAACCCAGCTGTGGCGGCATCATC 1269
 Db 381 TyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIle 400
 QY 1270 GGGAGAGGCGGCGACATCAACAGCTGGAGATTCGCCGAGCTCTTATCAAGATT 1329
 Db 401 GlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIle 420
 QY 1330 GCCCTCGGAAGCGCCAGACGTCAGCAAGAGTGTCTATCATCCGCGCCACCGGAA 1389
 Db 421 AlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGlu 440
 QY 1390 GCCAGTTCAGGCCAGGACCGATCTTTGGGAAACTGAAAGAGGAAATCTTTTAAAC 1449
 Db 441 AlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPheAsn 460
 QY 1450 CCGAAGAGCACTGAGCTGGAGCGCATATCAGAGTCCCTCTTCCAGCTGGCCGG 1509
 Db 461 ProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGlyArg 480
 QY 1510 GTGATTGGCAAGGTCGCAAGACCGTGAACGAACTGCAAGACTTAACCAAGTGCAGAGTC 1569
 Db 481 ValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal 500
 QY 1570 ATCGTCTCTGTCACCAAGCGCAGATGAAATGAGGAAGTATGCTCAGATTTATCGGG 1629
 Db 501 IleValProArgAspGlnThrProAspGluAsnGluValIleValArgIleLeGly 520
 QY 1630 CACTTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATTTGTAACAGAGTGAAG 1689
 Db 521 HisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGlnValLys 540
 QY 1690 CAGCAGAGCAGAAATACCTCAGGAGTGCCTCAGCAGCGCAAG 1737
 Db 541 GlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 3
 ID ABG06794
 XX ABG06794 standard; protein; 614 AA.
 AC ABG06794;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #6785.
 XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI N-PSDB; AAS70981.
 DR WPI; 2001-639362/73.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
 PT
 PT
 PT
 XX Claim 20; SEQ ID NO 37153; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Alignment Scores:
 Pred. No.: 0 Length: 614
 Score: 361.00 Matches: 361
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.52% Indels: 0
 DB: 4 Gaps: 0
 US-09-270-437D-8 (1-3283) x ABG06794 (1-614)

QY 58 GGGAGAGAGCGATGATGAACAGCTTACATCGGAGACCTGAGCCCGCTCACGGCC 117
 Db 12 GlyLysArgArgMetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAla 31
 QY 118 GAGACCTCGGAGCTCTTTGGGAGACAGAACTCCCTCGCGGACAGGTCTGCTG 177
 Db 32 AspAspLeuArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeu 51
 QY 178 AAGTCGGCTACGCTTCGTGGACTACCCGACAGAACTGGCCATCGCGCATCGAG 237
 Db 52 LysSerGlyTyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlu 71

QY 238 ACCCTCTCGGTAAGTGAATGCAATGCGGAAAATCATGGAAGTTGATTACTCAGTCCT 297
 Db 72 ThrLeuSerGlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSer 91
 QY 298 AAAAGCTAAGCAGCAGGAGAAATTCAGATCCAAATCCCTCCCTCAGTCCAGTGGAG 357
 Db 92 LysLysLeuArgSerArgLysIleGlnIleArgGlnIleProHisLeuGlnIleProGlu 111
 QY 358 GTGTTCGATGGACTTTTGGCTCAATATGGACAGTGGAGAAATGTGGAACTCAACACA 417
 Db 112 ValLeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThr 131
 QY 418 GACACAGAAACCGCCCTTCAACGTCACATATGCAACAGCAGAGCAACAAATAGCC 477
 Db 132 AspThrGluThrAlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAla 151
 QY 478 ATGGAGAAGCTAAGCGCGCATCAGTTTGAGAACTACTCTCTCAAGATTTCTACATCCCG 537
 Db 152 MetGluLysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIlePro 171
 QY 538 GATGAAGAGTGAAGTCCCTTGGCCCTCAGCGAGCCAGCGTGGAGCACTCTTCC 597
 Db 172 AspGluGluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSer 191
 QY 598 CGGAGCAGGCGCACCGCCCTGGGGCACTTCTCAGGCGCAGACAGATTGATTCCCGCTG 657
 Db 192 ArgGluGlnGlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeu 211
 QY 658 CGGATCCTGCTCCCAACCCAGCTTGTGTGGCCATCATCGGAAAGAGCGCTTGACCAT 717
 Db 212 ArgIleLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIle 231
 QY 718 AGAACAATCACTAAGCAGACCCAGTCCCGGTGATATCATAGAAAGAGCACTCTGA 777
 Db 232 LysAsnIleThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGly 251
 QY 778 GCTGCAGAGAGCTGTCCATCCATGCCAGAGGAGGACTTCTGAAGCATGCCGC 837
 Db 252 AlaAlaGluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArg 271
 QY 838 ATGATCTTGAATCATGCAAGAAGCAGCAGATGAGACCAATAGCCGAGAGATTCCT 897
 Db 272 MetIleLeuIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIlePro 291
 QY 898 CTGAATCTTGGCAGCAGATGCTGTGTGGAAGACTGATTCGAAAAGAGCAGCAAT 957
 Db 292 LeuLysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsn 311
 QY 958 TTGAAGAAAATTGAACATGAAACAGGAGCAGAGTAAATCTCATCTTTGCAAGATTG 1017
 Db 312 LeuLysLysIleGluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeu 331
 QY 1018 AGCATATACACCGGAGAACCATCTGTGAGGGCAGCAGTTCAGGCGCTGTGCGCAT 1077
 Db 332 SerIleTyrAsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSer 351
 QY 1078 GCTCAGATAGATATTGAAGAAGCTCGCTGAGCGCTTTGAAATGATATGCTGGCTGT 1137
 Db 352 AlaGluIleGluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaVal 371
 QY 1138 AAC 1140
 Db 372 Asn 372

RESULT 4

AAU16163

ID AAU16163 standard; protein; 620 AA.

XX AC

XX AAU16163;

XX AC

DT 07-NOV-2001 (first entry)

XX

Human novel secreted protein, Seq ID 1116.

Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
 cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;
 antibacterial; virucide; fungicide; opthalmological; vulnerary;
 secreted protein; rheumatoid arthritis; hyperproliferative disorder;
 cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 cerebral ischaemia; angiogenesis; nervous system disorder;
 Alzheimer's disease; infection; ocular disorder; corneal infection;
 wound healing; epithelial cell proliferation; skin ageing; food additive;
 preservative; antiproliferative.

Homo sapiens.

WO200155322-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001341.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209457P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216880P.

11-JUL-2000; 2000US-0217487P.

11-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225213P.

14-AUG-2000; 2000US-0225214P.

14-AUG-2000; 2000US-0225266P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

14-AUG-2000; 2000US-0225759P.

18-AUG-2000; 2000US-0226279P.

22-AUG-2000; 2000US-0226681P.

22-AUG-2000; 2000US-0226688P.

22-AUG-2000; 2000US-0227182P.

23-AUG-2000; 2000US-0227009P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229344P.

01-SEP-2000; 2000US-0229345P.

05-SEP-2000; 2000US-0229509P.

05-SEP-2000; 2000US-0229513P.

06-SEP-2000; 2000US-0230437P.

06-SEP-2000; 2000US-0230438P.

08-SEP-2000; 2000US-0231242P.

08-SEP-2000; 2000US-0231243P.

08-SEP-2000; 2000US-0231244P.

08-SEP-2000; 2000US-0231413P.

08-SEP-2000; 2000US-0231414P.

08-SEP-2000; 2000US-0232080P.

08-SEP-2000; 2000US-0232081P.

12-SEP-2000; 2000US-0231968P.

14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234987P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235837P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241212P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249266P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-489783/53.
 DR N-PSDB; AAS26150.
 XX New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX Claim 11; SEQ ID NO 1116; 980pp; English.
 PS The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Alignment Scores:
 Pred. No.: 0 Length: 620
 Score: 357.00 Matches: 357
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.15% Indels: 0
 DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x AAU16163 (1-620)

QY 70 ATGATGACAGCTTACATCGGAGACCTGAGCCGCCCTCCAGCCGACGACCTCCGG 129
 Db 22 MetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41
 QY 130 CAGCTCTTTGGGGACAGAAAGCTGCCCTCGCGGACAGGTCTCTGCTGAAGTCCGGTAC 189
 Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61
 QY 190 GCCTTCGTGATACCCCGACAGAACTGGGCCATCGGCCATCGAGACCTCTCGGGT 249
 Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81
 QY 250 AAAGTGAATGATCGGAAATCATGGAAGTGTATTACTCAGTCTCTAAAGCTAAGG 309

PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.

XX Claim 11; SEQ ID NO 1116; 402pp; English.

CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
 CC ABUS5748 represent human novel polypeptides of the invention

XX Sequence 620 AA;

Alignment Scores:

Pred. No.: 0 Length: 620
 Score: 357.00 Matches: 357
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.15% Indels: 0
 DB: 6 Gaps: 0

US-09-270-437D-8 (1-3283) x ABUS5232 (1-620)

QY 70 ATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCCGTCACCGCCGACCTCCGG 129
 DB 22 MetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41
 QY 130 CAGCTCTTTGGGACAGGAGCTCCCGTGGGACAGCTCTGCTGAAGTCCGCTAC 189
 DB 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61
 QY 190 GCCTTCGTGACTACCCGACACAGAACTGGGCCATCGGCCATCGAGACCTCTCGGT 249
 DB 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81
 QY 250 AAAGTGAANTGATGGGAATCATGGAATGATGATCTCAGTCTCTAAAAGCTAAG 309
 DB 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101
 QY 310 AGCAGGAATAATTCAGATTCTGAAACATCCCTCCTCAGCTCAGTGGGAGGTGGATGA 369
 DB 102 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 121
 QY 370 CTTTTCGCTCAATATGGAACAGTGGAGATGGAACAAAGTCAACACACACAGAAACC 429
 DB 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141
 QY 430 GCCGTGTGCAAGTCAATATGCAACAGAGAGAGCAAAATAGCCATGCGAGAGCTA 489
 DB 142 AlaValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161
 QY 490 AGCGGCATCAGTTTGAAGACTACTCTTCAAGATTCTACATCCCGATCAAGAGGTG 549
 DB 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 181
 QY 550 AGCTCCCTTCGCCCCCTCAGGAGCCAGCTGGGACCATCTCTCCCGGAGCAAGCC 609
 DB 182 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGln 201
 QY 610 CACGCCCTTCGGGGCAGCTTCTCAGGCCACAGAGATTGATTTCCCGCTCGGATCCTGTC 669
 DB 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221

QY 670 CCACCCAGTTTGTGTTGTCCTCATCCGAAAGGAGGCTTGCACCAATAAGACATCACT 729
 DB 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241
 QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGGAGCTGCAGAGAAG 789
 DB 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaGluLys 261
 QY 790 CTTGTCACATCCATCCGACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAA 849
 DB 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281
 QY 850 ATCATGCGAAAGAGCGCATGAGACCAAACTAGCCGAGAGATTCCTCTCAAAATCTTG 909
 DB 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 301
 QY 910 GCACACATGGCTTGGTGGAGACTGATGAGAAAGAGGAGCAAAATTTGAAGAAAT 969
 DB 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 321
 QY 970 GAACATGAACAGGAGCAACAGATCAATCTCATCTTTGACAGGATTTGAGCATATACAAC 1029
 DB 322 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 341
 QY 1030 CGGAAAGAACCATCATCTGTGAAGGCGACAGTTGAGGCTGTGCGAGTGTGAGATAGAG 1089
 DB 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361
 QY 1090 ATTATGAAGAGCTCGCTGAGGCTTTGAAAATGATATGCTGGCTGTAAAC 1140
 DB 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 378

RESULT 6

AAU16166
 ID AAU16166 standard; protein; 171 AA.
 AC AAU16166;
 XX 07-NOV-2001 (first entry)
 XX Human novel secreted protein, Seq ID 1119.
 DE Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
 KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
 KW cerebrovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;
 KW preservative; antiproliferative.
 XX Homo sapiens.
 OS WO200155322-A2.
 PN 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US0001341.
 PF 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216477P.
 PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-02117487P.
 PR 11-JUL-2000; 2000US-02117496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246605P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

N-PSDB; AAS26153.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 11; SEQ ID NO 1119; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune

CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and to maintain organ proliferation,
 CC transplanted tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Alignment Scores:
 Pred. No.: 1,36e-143 Length: 171
 Score: 160.00 Matches: 160
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.86% Indels: 0
 DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x AAU16166 (1-171)

Qy 1258 GGCGCCATCATCGGGAAGAGGGGGGACATCAACAGTGGCGAGATTCGGCGAGCC 1317
 Db 12 GlyAlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
 Qy 1318 TCTATCAAGATTGGCCCTCGGGAAGGCCAGAGCTCAGCGAAAGGATGTCATCATCAC 1377
 Db 32 SerIleLysIleAlaProIleGluGlyProAspValSerGluArgMetValIleIleThr 51
 Qy 1378 GGCGCCACCGAAGCCAGTTCAAGGCCCGAGCGAGTCTTCGGAACCTGAAAGAGGAA 1437
 Db 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLysLysGluGlu 71
 Qy 1438 AACTCTTTTACCCCAAGAGAGTGAAGCTGAGCGCCATATCAGAGTGCCTCTTCC 1497
 Db 72 AspPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
 Qy 1498 ACAGTCCGCGGTGATGTCGCAAGTGGCAAGCCGTCGAGCAAGTGCAGAACTTAAC 1557
 Db 92 ThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
 Qy 1558 AGTCGAGAGTATCGTCTGCTGCTGACCAACCCAGATGAAATCAGAGTGCATCGTC 1617
 Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
 Qy 1618 AGAATTATCGGCACCTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATTTGA 1677
 Db 132 ArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
 Qy 1678 CAACAGTGAAGCAGGAGGAGAGAAATACCTTCAGGGAGTGCCTCAGCGCAGCAAG 1737
 Db 152 GlnGlnValLysGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171

RESULT 7
 ABU55235
 ID ABU55235 standard; protein; 171 AA.
 XX AC ABU55235;
 XX DT 18-MAR-2003 (first entry)
 XX DE Human novel polypeptide #322.

XX Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;

KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 OS Homo sapiens.
 XX US2002132753-A1.
 XX 19-SEP-2002.
 XX 17-JAN-2001; 2001US-00764864.
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-023935P.
 PR 20-OCT-2000; 2000US-024060P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-147444/14.
 DR N-PSDB; ABX73494.
 XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,

PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.

Claim 11; SEQ ID NO 1119; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
XX ABUS5748 represent human novel polypeptides of the invention

SQ Sequence 171 AA;

Alignment Scores:

Pred. No.: 1.36e-143 Length: 171
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.86% Indels: 0
DB: 6 Gaps: 0

US-09-270-437D-8 (1-3283) x ABUS5235 (1-171)

QY 1258 GCGCCATCATCGGAGAGGGGCGACATCAACAGCTGGCGAGATTCGCGAGCC 1317
DB 12 GlyAlaIleGlyLysGlyAlaHisLysGlnLeuAlaArgPheAlaGlyAla 31
QY 1318 TCTATCAAGATTCGCTCGGAAAGCCCGAGAGCTGAGGAAAGGATGTCATCATACC 1377
DB 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51
QY 1378 GGGCCACCGAGCCCGATTCAGGCCCGAGGACGATCTTTGGGAAACTGGAAGGAA 1437
DB 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1438 AACTCTTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCC 1497
DB 72 AsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
QY 1498 ACAGTGGCCGGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTGAGAACTAAC 1557
DB 92 ThrAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
QY 1558 AGTGCAAGATCATCTGCTGCTGTCAGCAAGCCAGATGAAATGAGGAAGTGTGCTC 1617
DB 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
QY 1618 AGAATTATCGGCACATCTTTTGTACCCAGACTGCAACGCAAGATCAGGAATTTGTA 1677
DB 132 ArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1678 CAACAGGTGAAGCAGAGGAGCAAGATACCTCAGGAGTGCCTCAGACGCGCAGCAAG 1737
DB 152 GlnGlnValLysGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171

RESULT 8

AA030261

ID AAG03261 standard; protein; 97 AA.

XX AAG03261;

XX 06-OCT-2000 (first entry)

XX

DE

Human secreted protein, SEQ ID NO: 7342.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.

OS Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC03267.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 13; SEQ ID NO 7342; 71pp + Sequence Listing; English.

XX The present sequence is a polypeptide encoded by one of a large number of
XX 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
XX prepared from total human RNAs or polyA+ RNAs derived from 30 different
XX tissues. EST sequences usually correspond mainly to the 3' untranslated
XX region (UTR) of the mRNA because they are often obtained from oligo-dT
XX primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
XX sequences derived from the 5' ends of mRNAs and even in those cases where
XX longer cDNA sequences have been obtained, the full 5' UTR is rarely
XX included. 5' ESTs are derived from mRNAs with intact 5' ends and can
XX therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
XX are also used in diagnostic, forensic, gene therapy and chromosome
XX mapping procedures. They are used to obtain upstream regulatory sequences
XX and to design expression and secretion vectors

SQ Sequence 97 AA;

Alignment Scores:

Pred. No.: 3.05e-83 Length: 97
Score: 97.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.01% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x AAG03261 (1-97)

QY 70 ATGATGAACAGCTTTTACATCGGAACCTGAGCCCCGCCCTCACGCCGACGACCTCGG 129
DB 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTTGGGACAGGAAGCTGCCCTCGCGGACAGGCTCCTGTAAGTCCGGCTAC 189
DB 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCCGACAGAACTGGGCCCATCGCGCCATCGAGACCCCTCTCGGGT 249
DB 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAGTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGCTAAGG 309
DB 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTGGAACATCCCTCCTCAGCTGCAAGTGGAGGTG 360
DB 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 97

RESULT 9
AAU16164
ID AAU16164 standard; protein; 192 AA.
XX AC AAU16164;
XX DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 1117.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmologic; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX PN WO200155322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001341.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225514P.
PR 14-AUG-2000; 2000US-0225466P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251989P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26151.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1117; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemoraxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Alignment Scores:
Pred. No.: 1,97e-57 Length: 192
Score: 70.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.50% Indels: 0
DB: 4 Gaps: 0
US-09-270-437D-8 (1-3283) x AAU16164 (1-192)
QY 781 GCAGAGAGGCTGTACCATTCATGCGCCAGAGGGGACTTCTGAGCATGCCCATG 840
DB 1 AlaGluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMet 20
QY 841 ATCTTGAAATCATGCGAGAGAGGAGGAGATGAGACCAACTAGCCGAGAGAGATTCCTCTG 900

Db 21 IleLeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeu 40
QY 901 AAAATCTTGCCACACAAATGCTTGGTTGGAAGACTGATTGGAAGAGAGGAGCAAAATTG 960
Db 41 LysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 60
QY 961 AAGAAATTCACATCAACACGAGGACCAAG 990
Db 61 LysLysIleGluHisGluThrGlyThrLys 70
RESULT 10
ABUS5233
ID ABUS5233 standard; protein; 192 AA.
XX
XX AC ABUS5233;
XX
XX DT 18-MAR-2003 (first entry)
XX
XX DE Human novel polypeptide #320.
XX
XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX OS Homo sapiens.
XX
XX PN US2002132753-A1.
XX
XX PD 19-SEP-2002.
XX
XX PP 17-JAN-2001; 2001US-00764864.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73492.
XX
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 11; SEQ ID NO 1117; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 192 AA;

Alignment Scores:
Pred. No.: 1,97e-57 Length: 192
Score: 70.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.50% Indels: 0
DB: Gaps: 0

US-09-270-437D-8 (1-3283) x ABU55233 (1-192)

QY 781 GCAGAGAGCCTGTCCACATCCATCCAGCCAGAGGGAGCTTCTGAAGCATGCCGATG 840
Db 1 AAGLuLeuProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgWet 20

QY 841 ATTCTTGAATCATGCAGAAAGAGGCAGATGAGCAAACTAGCCGAAAGAGATTCTCTG 900
Db 21 IleLeuGluLeuMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeu 40

QY 901 AAAATCTTGGCACACATGGCTTGGTTGAAGACTGATGGAAAGAGCGGAGAAATTG 960
Db 41 LysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 60

QY 961 AAGAAATTTGAATCAATGAACAGGACCAAG 990
|||||

Db 61 LysLysIleGluHisGluThrGlyThrLys 70
RESULT 11
AAU16583
ID AAU16583 standard; protein; 171 AA.
XX
XX AAU16583;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human novel secreted protein, Seq ID 1536.
XX
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.


```

Db      146 LysileArgGluIleVal 151
RESULT 13
ABU89799
ID      ABU89799 standard; protein; 555 AA.
XX
XX
AC      AC
DT      10-JUL-2003 (first entry)
XX
DE      Novel human protein NOV14a.
XX
KW      Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.
XX
OS      Homo sapiens.
XX
PN      WO2003031571-A2.
XX
PD      17-APR-2003.
XX
PF      02-OCT-2002; 2002WO-US0311357.
XX
PR      05-OCT-2001; 2001US-0327454P.
PR      09-OCT-2001; 2001US-0327917P.
PR      09-OCT-2001; 2001US-0328029P.
PR      09-OCT-2001; 2001US-0328056P.
PR      12-OCT-2001; 2001US-0328849P.
PR      15-OCT-2001; 2001US-0328414P.
PR      17-OCT-2001; 2001US-0330142P.
PR      22-OCT-2001; 2001US-0341058P.
PR      24-OCT-2001; 2001US-0343629P.
PR      29-OCT-2001; 2001US-0349575P.
PR      01-NOV-2001; 2001US-0346357P.
PR      25-JUN-2002; 2002US-0391342P.
PR      01-OCT-2002; 2002US-00262445.
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Alcobrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;
PI      Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;
PI      Mezes PS, Millet I, Ooi CE, Patturajan M, Rieger DK, Spytek KA;
PI      Taupier RJ, Zethusen BD, Zhong H, Zhong M;
XX      WPI; 2003-381704/36.
DR      N-PSDB; ACA90176.
XX
XX
PT      New DAPK3 polypeptide, useful for preparing a composition for treating or
PT      preventing e.g., cancer.
XX
PS      Claim 2; Page 129; 253pp; English.
XX
CC      The invention describes an isolated polypeptide comprising any of 33 90-
CC      1273 amino acid sequences (I) given in the specification or its mature
CC      form, a sequence that is at least 95 % identical to (I), or a sequence
CC      comprising one or more conservative substitutions in the amino acid
CC      sequence of (I). The polypeptide is useful for preparing a composition
CC      for treating or preventing e.g. cancer. This is the amino acid sequence
CC      of a novel human NOV protein
XX
SQ      Sequence 555 AA;

Alignment Scores:
Pred. No.:      1.32e-32      Length:      555
Score:          44.00      Matches:      44
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      4.09%      Indels:      0
DB:              6      Gaps:        0

US-09-270-437D-8 (1-3283) x ABU89799 (1-555)
QY      325 ATTGGAACATCCCTCCTCAGTGGAGGTGTGTGATGGACTTTTGGCTCAATAT 384

```

```

Db      86 IleArgGlnIleProHisLeuGlnTrpGluValLeuAspGlyLeuLeuAlaGlnTyr 105
QY      385 GGGACACTGGAGATGTGGACAGTCAACACAGACACAGACAGACCGCTTGTCACTC 444
Db      106 GlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValValAsnVal 125
QY      445 ACATATGCAACA 456
Db      126 ThrTyrAlaThr 129

RESULT 14
AAO23971
ID      AAO23971 standard; protein; 187 AA.
XX
XX      AAO23971;
XX
XX      06-NOV-2003 (first entry)
XX
XX      Human IGF-II mRNA binding protein 3.
XX
KW      Anorectic; antiinflammatory; cardiant; hypotensive; antidiabetic;
KW      neuroprotective; pharmaceutical composition; body-weight regulation;
KW      thermogenesis; metabolic; obesity; Syndrome X; insulin-resistance;
KW      eating disorder; cachexia; diabetes mellitus; hypertension; gallstone;
KW      pancreatic dysfunction; arteriosclerosis; coronary heart disease;
KW      hypercholesterolaemia; dyslipidaemia; osteoarthritis; ROS defence;
KW      reactive oxygen species; neurodegenerative; mitochondrial; gene therapy;
KW      human; IGF-II mRNA binding protein 3.
XX
OS      Homo sapiens.
XX
PN      WO2003061681-A2.
XX
PD      31-JUL-2003.
XX
XX      24-JAN-2003; 2003WO-EP000738.
XX
XX      25-JAN-2002; 2002EP-00001806.
PR      14-FEB-2002; 2002EP-00003473.
PR      28-FEB-2002; 2002EP-00004687.
PR      25-APR-2002; 2002EP-00009475.
PR      18-JUN-2002; 2002EP-00013329.
PR      30-DEC-2002; 2002EP-00029081.
XX
XX      (DEVE-) DEVELOPENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
XX      Steuernagel A, Molitor A, Eulenberg K, Broenner G;
XX      WPI; 2003-627418/59.
XX      N-PSDB; AAL57525.
XX
XX      New pharmaceutical composition, useful for the manufacture of an agent
XX      for diagnosing, treating or preventing disorders related to body-weight
XX      regulation and thermogenesis, e.g., metabolic diseases such as obesity.
XX
XX      Claim 3; Fig 7C; 144pp; English.
XX
CC      The invention relates to a novel pharmaceutical composition comprising a
CC      nucleic acid molecule or polypeptide which is a human homologue of a
CC      Drosophila melanogaster polypeptide or polynucleotide. The composition of
CC      the invention may be utilised during the diagnosis, study, prevention and
CC      treatment of diseases related to body-weight regulation and thermogenesis
CC      including metabolic disorders such as obesity, Syndrome X and insulin-
CC      resistance syndrome and eating disorders e.g. cachexia, diabetes
CC      mellitus, hypertension, pancreatic dysfunctions, arteriosclerosis,
CC      coronary heart disease, hypercholesterolaemia, dyslipidaemia,
CC      osteoarthritis and gallstones. Furthermore, disorders related to reactive
CC      oxygen species (ROS) defence may be addressed by the invention including
CC      neurodegenerative disorders or mitochondrial disorders. Finally, the
CC      composition of the invention may be useful in gene therapy. The current
CC      sequence is that of the human IGF-II mRNA binding protein 3 protein of
CC      the invention

```

```
XX SQ Sequence 187 AA;
Alignment Scores:
Pred. No.: 4.76e-21 Length: 187
Score: 32.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 6 Gaps: 0

US-09-270-437D-8 (1-3283) x AA023971 (1-187)

QY 1381 CCACCGAAGCCCAAGTTCAAGGCCAGGACGGATCTTTGGAAACTGAAAGAGGAAAC 1440
Db 97 ProProGluAlaGlnPheLysAlaGlnGlyArgilePheGlyLysLeuLysGluGluAsn 116

QY 1441 TTCTTTAACCACCAAGAGAGAGTGAAGTGAAGCG 1476
Db 117 PhePheAsnProLysGluGluValLysLeuGluAla 128

RESULT 15
AAU16581
ID AAU16581 standard; protein; 47 AA.
AC AAU16581;
XX
XX 07-NOV-2001 (first entry)
XX Human novel secreted protein, Seq ID 1534.
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulneryary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX Homo sapiens.
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-0235484P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 02-OCT-2000; 2000US-0236370P.
XX 02-OCT-2000; 2000US-0236802P.
XX 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
XX 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
XX 13-OCT-2000; 2000US-0239937P.
XX 13-OCT-2000; 2000US-0239935P.
XX 20-OCT-2000; 2000US-0240960P.
XX 20-OCT-2000; 2000US-0241221P.
XX 20-OCT-2000; 2000US-0241785P.
XX 20-OCT-2000; 2000US-0241786P.
XX 20-OCT-2000; 2000US-0241787P.
XX 20-OCT-2000; 2000US-0241808P.
XX 20-OCT-2000; 2000US-0241809P.
XX 01-NOV-2000; 2000US-0241826P.
XX 08-NOV-2000; 2000US-0244517P.
XX 08-NOV-2000; 2000US-0244574P.
XX 08-NOV-2000; 2000US-0246475P.
XX 08-NOV-2000; 2000US-0246476P.
XX 08-NOV-2000; 2000US-0246477P.
XX 08-NOV-2000; 2000US-0246478P.
XX 08-NOV-2000; 2000US-0246523P.
XX 08-NOV-2000; 2000US-0246524P.
XX 08-NOV-2000; 2000US-0248525P.
XX 08-NOV-2000; 2000US-0248526P.
```

PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26568.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1534; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
XX encoded secreted proteins. The nucleic acids and proteins are used to
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
XX in diagnosing a pathological condition or susceptibility to a
XX pathological condition. Antibodies to the proteins can also be used in
XX alleviating symptoms associated with the disorders and in diagnostic
XX immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
XX (ELISA). Disorders which are diagnosed or treated include autoimmune
XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
XX neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX nervous system disorders e.g. Alzheimer's disease, infections caused by
XX bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
XX and many other disorders listed in the specification. The polypeptides
XX can also be used to aid wound healing and epithelial cell proliferation,
XX to prevent skin aging due to sunburn, to maintain organs before
XX transplantation, for supporting cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage

CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Alignment Scores:
Pred. No.: 1.78e-09 Length: 47
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.86% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x AAU16581 (1-47)

Qy 643 ATTGATTTCCGCTGCGGATCCCTGTCCTCCACCCAGTTTCTTGGTGCATCATCGGAAG 702
Db 1 IleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIleIleGlyLys 20

Search completed: August 6, 2004, 13:20:43
Job time : 158.09 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:27 ; Search time 29.5087 Seconds
(without alignments)
11487.328 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcagcgagagcgagga.....aaccttgaaatgtttattt 3283

Scoring table:

Xgapop	60.0	Xgapext	60.0
Ygapop	60.0	Ygapext	60.0
Zgapop	6.0	Zgapext	7.0
Delop	6.0	Delext	7.0

Searched: 389414 seqs, 51625971 residues
Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/uspto.spool_p/US09270437/runat_06082004_141221_1858/app_query.fasta_1.5582
-DB=Issued Patents AA -OFFIT=fastan -SUFFIX=oligo806.ra -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -HEAPS=500 -MINLEN=0
-ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPS=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US09270437 @CGN 1.1.59 @runat_06082004_141221_1858
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6C_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/6D_COMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	1.8	48	3	US-09-261-855-20
2	19	1.8	577	3	US-09-261-855-2
3	19	1.8	579	4	US-09-643-597-176
4	19	1.8	579	4	US-09-643-597-348
5	19	1.8	579	4	US-09-480-884A-176
6	19	1.8	579	4	US-09-542-615A-176
7	19	1.8	579	4	US-09-542-615A-348
8	19	1.8	579	4	US-09-606-421B-176
9	19	1.8	579	4	US-09-606-421B-348
10	18	1.7	48	3	US-09-261-855-24
11	16	1.5	49	3	US-09-261-855-22
12	14	1.3	49	3	US-09-261-855-18

13	13	1-2	47	3	US-09-261-855-23	Sequence 21, Appl
14	13	1-2	47	3	US-09-261-855-21	Sequence 23, Appl
15	11	1-0	47	3	US-09-261-855-19	Sequence 19, Appl
c 16	9	0-8	30	2	US-08-723-306-25	Sequence 25, Appl
c 17	9	0-8	30	2	US-08-723-306-30	Sequence 30, Appl
c 18	9	0-8	30	5	PCT-US96-10041-25	Sequence 25, Appl
c 19	9	0-8	30	5	PCT-US96-10041-30	Sequence 30, Appl
c 20	9	0-8	47	3	US-09-261-855-17	Sequence 17, Appl
c 21	9	0-8	81	4	US-09-489-039A-8805	Sequence 8805, Ap
c 22	9	0-8	399	4	US-09-252-991A-31280	Sequence 31280, A
c 23	9	0-8	491	1	US-08-206-176-4	Sequence 4, Appl
c 24	9	0-8	603	4	US-09-543-681A-6975	Sequence 6975, Ap
c 25	9	0-8	657	4	US-09-252-991A-28001	Sequence 28001, A
26	8	0-7	11	3	US-09-261-855-11	Sequence 11, Appl
27	8	0-7	11	3	US-09-261-855-13	Sequence 13, Appl
c 28	8	0-8	70	4	US-09-134-001C-4634	Sequence 4634, Ap
c 29	8	0-8	74	4	US-09-621-976-6160	Sequence 6160, Ap
c 30	8	0-8	117	4	US-09-134-001C-5606	Sequence 5606, Ap
c 31	8	0-8	136	4	US-09-252-991A-26248	Sequence 26248, A
c 32	8	0-7	171	2	US-08-560-098A-53	Sequence 53, Appl
c 33	8	0-8	175	4	US-09-252-991A-31050	Sequence 31050, A
c 34	8	0-7	204	4	US-09-489-039A-7570	Sequence 7570, Ap
c 35	8	0-8	214	4	US-09-489-847-169	Sequence 169, App
c 36	8	0-8	223	4	US-09-724-864-42	Sequence 42, Appl
c 37	8	0-8	258	4	US-09-252-991A-21821	Sequence 21821, A
c 38	8	0-8	258	4	US-09-252-991A-27264	Sequence 27264, A
c 39	8	0-7	260	4	US-09-540-236-2747	Sequence 2747, Ap
c 40	8	0-7	264	4	US-09-252-991A-21057	Sequence 21057, A
c 41	8	0-7	265	4	US-09-252-991A-17835	Sequence 17835, A
c 42	8	0-7	279	4	US-09-252-991A-31036	Sequence 31036, A
c 43	8	0-8	290	4	US-09-489-847-327	Sequence 327, App
c 44	8	0-8	322	4	US-09-134-000C-5510	Sequence 5510, Ap
c 45	8	0-7	332	4	US-09-252-991A-21597	Sequence 21597, A

ALIGNMENTS

RESULT 1
US-09-261-855-20
; Sequence 20, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261.855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-20

Alignment Scores:	7,75e-09	Length:	48
Pred. No.:	19.00	Matches:	19
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	1.76%	Indels:	0
Query Match:	3	Gaps:	0
DB:			

US-09-270-437D-8 (1-3283) x US-09-261-855-20 (1-48)

QY 1501 GTGTCCCGGTGTAATGGCAAGGTGGCAAGACCGTGAACGACTGCAGACTTACC 1557
Db 8 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 26

RESULT 2
US-09-261-855-2
; Sequence 2, Application US/09261855A

```

RESULT 4
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Panger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 389
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.:      4,93e-09      Length:      579
Score:          19.00        Matches:     19
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     1.76%       Indels:      0
DB:              4          Gaps:        0

US-09-270-437D-8 (1-3283) x US-09-643-597-348 (1-579)

QY      1360 AGGATGGTCATCATCCGGCCACCGGAAGCCAGTTCACGGCCAGGACGGATC 1416
Db      452 ArgMetValIleIleThrGlyProProGluAlaGlnPheIysAlaGlnGlyArgIle 470

RESULT 5
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Panger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:      4,93e-09      Length:      579
Score:          19.00        Matches:     19
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     1.76%       Indels:      0
DB:              4          Gaps:        0

```


US-09-270-437D-8 (1-3283) x US-09-480-884A-176 (1-579)

QY 1360 AGGATGTCATCATCACCGGCGCCACCGAGCCAGTTCAGGCCCGGAGCGGATC 1416
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 6

US-09-542-615A-176
; Sequence 176, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fang, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542.615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-542-615A-176

Alignment Scores:

Pred. No.: 4.93e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-542-615A-176 (1-579)

QY 1360 AGGATGTCATCATCACCGGCGCCACCGAGCCAGTTCAGGCCCGGAGCGGATC 1416
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 7

US-09-542-615A-348

; Sequence 348, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fang, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542.615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 348

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-542-615A-348

Alignment Scores:

Pred. No.: 4.93e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-542-615A-348 (1-579)

QY 1360 AGGATGTCATCATCACCGGCGCCACCGAGCCAGTTCAGGCCCGGAGCGGATC 1416
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 8

US-09-606-421B-176

; Sequence 176, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy R.

; APPLICANT: Fang, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606.421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-606-421B-176

Alignment Scores:

Pred. No.: 4.93e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-606-421B-176 (1-579)

QY 1360 AGGATGTCATCATCACCGGCGCCACCGAGCCAGTTCAGGCCCGGAGCGGATC 1416
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 9

US-09-606-421B-348

; Sequence 348, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fang, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606.421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 348

```
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 4.93e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-606-421B-348 (1-579)

QY 1360 AGGATGTCATCATCACCAGCCAGCCAGTTCAGGCCAGGACGATC 1416
      |||||
Db 452 ArgMetValIleIleThrGlyProProGluAlaGlnPheIysalaGlnGlyArgIle 470

RESULT 10
US-09-261-855-24
; Sequence 24, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-24

Alignment Scores:
Pred. No.: 6.58e-08 Length: 48
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.67% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-24 (1-48)

QY 1501 GCTGGCCGGTGTGTTGGCAAGTGGCAAGCCGTGAACGAACTGCAGAACTTA 1554
      |||||
Db 8 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeu 25

RESULT 11
US-09-261-855-22
; Sequence 22, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-22

Alignment Scores:
Pred. No.: 4.72e-06 Length: 49
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-22 (1-47)

; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 4.93e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-606-421B-348 (1-579)

QY 1360 AGGATGTCATCATCACCAGCCAGCCAGTTCAGGCCAGGACGATC 1416
      |||||
Db 452 ArgMetValIleIleThrGlyProProGluAlaGlnPheIysalaGlnGlyArgIle 470

RESULT 10
US-09-261-855-24
; Sequence 24, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-24

Alignment Scores:
Pred. No.: 6.58e-08 Length: 48
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.67% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-24 (1-48)

QY 1501 GCTGGCCGGTGTGTTGGCAAGTGGCAAGCCGTGAACGAACTGCAGAACTTA 1554
      |||||
Db 8 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeu 25

RESULT 11
US-09-261-855-22
; Sequence 22, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-22

Alignment Scores:
Pred. No.: 4.72e-06 Length: 49
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-22 (1-47)

Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.49% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-22 (1-49)

QY 925 GTTGAAGACTGATTCGAAAGAGCGCAGAAATTTGAAGAAATTCGAA 972
      |||||
Db 8 ValGlyArgLeuIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 23

RESULT 12
US-09-261-855-18
; Sequence 18, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-18

Alignment Scores:
Pred. No.: 0.00034 Length: 49
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-18 (1-49)

QY 925 GTTGAAGACTGATTCGAAAGAGCGCAGAAATTTGAAGAAATTCGAGAA 966
      |||||
Db 8 ValGlyArgLeuIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 21

RESULT 13
US-09-261-855-21
; Sequence 21, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-21

Alignment Scores:
Pred. No.: 0.00291 Length: 47
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-21 (1-47)
```

Job time : 40.5087 secs

QY 664 CTGCTCCACACACCTTGTGGTCCATCATCGGAAG 702
DB 2 LeuValProThrGlnPheValGlyAlaIleIleGlyLys 14

RESULT 14
US-09-261-855-23
; Sequence 23, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-23

Alignment Scores:
Pred. No.: 0.00291 Length: 47
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-23 (1-47)
QY 1303 AGATTCGGCGGACCTTATCAAGATTGCCCTGCGGAA 1341
DB 24 ArgPheAlaGlyAlaSerIleLysIleAlaProAlaGlu 36

RESULT 15
US-09-261-855-19
; Sequence 19, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-19

Alignment Scores:
Pred. No.: 0.21 Length: 47
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.02% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-261-855-19 (1-47)
QY 1249 CAGGCTGTGGCGGCATCATCGGAAGAAGGG 1281
DB 6 GlnAlaValGlyAlaIleIleGlyLysGly 16

Search completed: August 6, 2004, 13:35:26

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:31:07 ; Search time 157.275 Seconds
(without alignments)
13095.793 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcacggaggagcggagga.....aaccttgaaatgtttattt 3283

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1291235 seqs, 313682936 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2479628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool_p/US03270437/runat_06082004_141223_1936/app_query.fasta_1.5582
-DB=Published Applications AA -QPMT=fastan -SUFFIX=oligo806.rapp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US03270437@cgn_1.1.35 @runat_06082004_141223_1936
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pcp:
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pcp:
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pcp:
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pcp:
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pcp:
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pcp:
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pcp:
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pcp:
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pcp:
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pcp:
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pcp:
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pcp:
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pcp:
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pcp:
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pcp:
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pcp:
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pcp:
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-10-097-340-147
; Sequence 147, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLIS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030

Sequence 147, App
Sequence 182, App
Sequence 1116, Ap
Sequence 501, App
Sequence 1117, Ap
Sequence 1117, Ap
Sequence 1536, Ap
Sequence 40, Appl
Sequence 499, App
Sequence 1534, Ap
Sequence 20, Appl
Sequence 476, App
Sequence 476, App
Sequence 1532, Ap
Sequence 1114, Ap
Sequence 2088, Ap
Sequence 2, Appli
Sequence 500, App
Sequence 176, App
Sequence 348, App
Sequence 176, App
Sequence 348, App
Sequence 176, App
Sequence 348, App
Sequence 446, App
Sequence 176, App
Sequence 446, App
Sequence 446, App
Sequence 449, App
Sequence 480, App
Sequence 484, App
Sequence 176, App
Sequence 348, App
Sequence 446, App
Sequence 449, App
Sequence 480, App
Sequence 484, App
Sequence 427, App
Sequence 427, App

; CURRENT APPLICATION NUMBER: US/10/097,340
 ; CURRENT FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 60/276,025
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/325,149
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/276,026
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/324,967
 ; PRIOR FILING DATE: 2001/09/26
 ; PRIOR APPLICATION NUMBER: 60/311,732
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/325,102
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/323,580
 ; PRIOR FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 147
 ; LENGTH: 556
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-097-340-147

Alignment Scores:

Pred. No.: 0 Length: 556
 Score: 556.00 Matches: 556
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.62% Indels: 0
 DB: 14 Gaps: 0

US-09-270-437D-8 (1-3283) x US-10-097-340-147 (1-556)

QY 70 ATGATGAAAGCTTTTACATCGGAACCTGAGCCCGCTGACCGCGACGACCTCCGG 129
 Db 1 MetMetAsnLysLeuYrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
 QY 130 CAGCTCTTTGGGACAGGAGCTGCCCTGGCGGACAGTCTCTGAGTCCCGGTAC 189
 Db 21 GlnLeuPheIYAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 40
 QY 190 GCCTTCGTGACTACCCGACCACTGGGCGCATCCGCGCATCGAGACCTCTCCGGT 249
 Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
 QY 250 AAGTCGAATTCGATGGGAATCATCGAAGTTGATTACTAGTCTCTAAAAGCTAAG 309
 Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLeuArg 80
 QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGTGGGAGGTGTGATGGA 369
 Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
 QY 370 CTTTGGCTCAATATCGGACAGTGGAGATGTGAAACAAAGTCAACACAGACAGAAC 429
 Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
 QY 430 GCCTTGTCAAGCTCACATATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
 Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
 QY 490 ACGGGCATCAGTTTGAAGACTTCTTCAAGATTTCTTCAATCCCGATCCGAGAGGTG 549
 Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
 QY 550 AGCTCCCTTCGCCCTCAGGAGCCGAGCTGGGGACCTCTTCCCGGACGACAGG 609
 Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
 QY 610 CACGCCCTTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTCGGATCCTGTC 669
 Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200

QY 670 CCCACCCAGTTTGTGGTGCCATCATCGAAAGAGGAGCTTGACCAATAAGAACATCACT 729
 Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluLeuThrIleLysAsnIleThr 220
 QY 730 AAGCAGACCCAGTCCCGGTAGATATCATATAAGAGAAAGAACTCTGAGCTGCAGAGAAG 789
 Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
 QY 790 CCTGTCCATCCATGCCACCCAGAGGGGACTTCTCAAGCATGCCCATCATCTTCTGAA 849
 Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
 QY 850 ATCATGAGAAAGAGGAGCATGACACCAAACTAGCCGAGAGATTCCTCTGAAATCTTG 909
 Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
 QY 910 GCACACATGGCTTGGTGGAGACTGATTCGAAAGAGAGAGAGAGAGAGAGAGAGAT 969
 Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
 QY 970 CAACATGAAAACAGGACCAAGATAAACAATCTCTTTCAGGATTTGACATATACAAC 1029
 Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
 QY 1030 CCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGCCCTGTGCCAGTGTGAGATAGAG 1089
 Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
 QY 1090 ATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGCTGTGTACACCCACTCC 1149
 Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnThrHisSer 360
 QY 1150 GGATATCTTCCAGCTGTACCCCATCACCACTGGCCCTTCCCGCATCACTCT 1209
 Db 361 GlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHisHisHisSer 380
 QY 1210 TATCCAGACAGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGGCCATCANC 1269
 Db 381 TyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIle 400
 QY 1270 GGGAGAGAGGGGGCACATCAACAGCTGGCGAGATTCCCGGAGCCTCTATCAAGATT 1329
 Db 401 GlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIle 420
 QY 1330 GCCCTCCGGAAGGCCAGAGCTCAGGAAAGAGATGTCATCATCACCAGGCGCCACCGGAA 1389
 Db 421 AlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProGlu 440
 QY 1390 GCCCAGTTCAAGGCCAGGAGCGGATCTTTGGGAAACTGAAAGAGGAAAATTTCTTTAAC 1449
 Db 441 AlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPheAsn 460
 QY 1450 CCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTCCACAGCTGCCCGG 1509
 Db 461 ProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGlyArg 480
 QY 1510 GTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCAGAACTTAACACAGTGCAGAGTC 1569
 Db 481 ValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal 500
 QY 1570 ATCGTGCCTGTGACCAAGCGCAGATGAAATGAGGAAGTGAATCGTCAGATTTATCGGG 1629
 Db 501 IleValProArgAspGlnThrProAspGluAsnGluValIleValArgIleIleGly 520
 QY 1630 CACTCTTTCTAGCAGACTGCACAGCCAGATCAGGGAATTTGTACAACAGGTGAAG 1689
 Db 521 HisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGlnValLys 540
 QY 1690 CACGAGGACAGAAATVACCTTCAGGAGTCCGCTCACAGCGCAGCAAG 1737
 Db 541 GlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 2

US-10-648-593-182
 ; Sequence 182 Application US/10648593
 ; Publication No. US2004010632A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
 ; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
 ; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
 ; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
 ; FILE REFERENCE: D0273 NP
 ; CURRENT APPLICATION NUMBER: US/10/648,593
 ; CURRENT FILING DATE: 2003-08-26
 ; PRIOR APPLICATION NUMBER: 60/406,385
 ; PRIOR FILING DATE: 2002-08-27
 ; NUMBER OF SEQ ID NOS: 557
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 182
 ; LENGTH: 556
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-648-593-182

Alignment Scores:
 Pred. No.: 0 Length: 556
 Score: 556.00 Matches: 556
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.62% Indels: 0
 DB: 16 Gaps: 0

US-09-270-437D-8 (1-3283) x US-10-648-593-182 (1-556)

QY	70	ATGATGAAACAGCTTTACATCGGAAACCTGAGCCCGCGTCCACCGCCGACCTCCGG	129
Db	1	MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg	20
QY	130	CAGCTCTTTGGGACAGGAAGTGCCTCGCGGACAGCTCTCTGCTGAAGTCCGGCTAC	189
Db	21	GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr	40
QY	190	GCCTTCGTGACTACCCACACAGAACTGGGCATCCCGCCCATCAGAGACCTCTCGGT	249
Db	41	AlaPheValAspTyrProAspGlnMetTrpAlaIleArgAlaIleGluThrLeuSerGly	60
QY	250	AAAGTGAATTCATCGGAAATCATGGAAGTTGATTACTAGTCTCTAAAGCTAAGG	309
Db	61	LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg	80
QY	310	AGCAGGAAATTCAGATTGGAATCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	369
Db	81	SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly	100
QY	370	CTTTTGGCTCAATATGGACAGTGGAGATGTGGAACAGTCAACACACACAGAAACC	429
Db	101	LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr	120
QY	430	GCCTGTGTCACTCATATGCAACAGAGAAAGAGCAAAATAGCCATGAGAGACTA	489
Db	121	AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu	140
QY	490	AGCGGGCATCAGTTTGAGAACTACTCTCTCAAGATTCTTACATCCCGGATGAGAGT	549
Db	141	SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal	160
QY	550	AGTCCCTTCGCGCTCAGCAGCCCGCGTGGGACCACTCTTCCCGGAGCAGGCG	609
Db	161	SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly	180
QY	610	CAGCGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGTGGGATCTCGTC	669
Db	181	HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal	200
QY	670	CCACCCAGCTTTGTTGTGTCATCATCGGAAAGGAGGCTTGACCATAAAGACATCACT	729

RESULT 3

US-09-764-864-1116

Db	201	ProThrGlnPheValGlyAlaIleIleGlyLeuGlyLeuThrIleLysAsnIleThr	220
QY	730	AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAAGAGAACTCTCGAGCTGCAGAGAAG	789
Db	221	LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaGluLys	240
QY	790	CTGTGTCACCATCCATCCACCCAGAGGGAGCTCTTGAAGCATGCCGCTATGTTCTGAA	849
Db	241	ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu	260
QY	850	ATCATGCAGAAAGAGGAGAGATGAGCAAACTAGCCAAAGAGATTCCTCTGAAAATCTTG	909
Db	261	IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeu	280
QY	910	GCACACATCGCTTGGTTGGAAGACTGATTGGAAAAGAGAGGAGAAATTTGAAGAAAT	969
Db	281	AlaHisAsnGlyLeuValGlyArgLeuIleGlyGlyArgAsnLeuLysLysIle	300
QY	970	GACATGAAACAGGACCAAGATAAATCTCATCTTTGAGGATTTGACATATACAAC	1029
Db	301	GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn	320
QY	1030	CCGAAAAGAACCATCACTGTGAAGGGACAGTTGAGCCCTGTGCCAGTGTGATAGAG	1089
Db	321	ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu	340
QY	1090	ATTATGAAGAGCTGCGTGGAGGCTTTGAAAATGATATGCTGTGTTAAACACCCATCC	1149
Db	341	IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnThrHisSer	360
QY	1150	GGATPACTTCTCCAGCTGTATACCCCATCACAGTTTGGCCGCTTCCGCGATCATCTCT	1209
Db	361	GlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHisSer	380
QY	1210	TATCAGACAGCAGAGATGTGAATCTCTTCATCCACCCAGGCTGTGGCGCATCATC	1269
Db	381	TyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValAlaIleIle	400
QY	1270	GGGAAGAGGGGGGCACACATCAAAACAGCTGGCAGATTCGCCGAGGCTCTTATCAAGATT	1329
Db	401	GlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIle	420
QY	1330	CCCTCTCGGAGGCCAGCAGCTCAGGAAAGGATGTGTATCATCATCAGCGGCCACCGAA	1389
Db	421	AlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGlu	440
QY	1390	CCCGAGTTCAAGGCCAGGACCGATCTTTGGGAACTGAAAGAGAGAAAATCTTTTAAAC	1449
Db	441	AlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPheAsn	460
QY	1450	CCCAAAGAGAAGTCAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCAGAGTGGCGCG	1509
Db	461	ProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGlyArg	480
QY	1510	GTGATTGCAAGGTGCGCAAGACCGTGAACCACTGCAGAACTTAACAGTGCAGAACTC	1569
Db	481	ValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal	500
QY	1570	ATCGTGCCTCTGACAAAACCGCAGATGAAAATGAGGAAGTGCATCGTCAGAAATTTATCGG	1629
Db	501	IleValProArgAspGlnThrProAspGluAsnGluValIleValArgIleIleGly	520
QY	1630	CACCTCTTTGTCAGCAGCTGCAGCGCAAGTACAGGAATTTGTAACACAGTGAAG	1689
Db	521	HisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGlnValLys	540
QY	1690	CAGCAGGAGCAGAAATACCTCAGGAGTGCCTCAGCAGCGCAGCAAG	1737
Db	541	GlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys	556

```
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Alignment Scores:
Pred. No.: 3,09e-315 Length: 620
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 33.15% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-764-864-1116 (1-620)

QY 70 ATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGTCCACCGCCGACGACCTCGG 129
DB 22 MetMetAsnLysLeuThrGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41
QY 130 CAGCTCTTTGGGACAGGAGCTGCCCTGGCGGACAGCTCCTGCTGAAGTCCGGCTAC 189
DB 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 61
QY 190 GCCTTCGTGACTACCCGACGAGACTGGGCATCGGCGCATCGGACCTCTCGGCT 249
DB 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlnThrLeuSerGly 81
QY 250 AAAGTGAATTCATGGGAAATCACTGGAAGTTGATTCTCACTCTCTATAAAGCTAAGG 309
DB 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101
QY 310 AGCAGGAAATTCAGATTGAAACATCCCTCCTACCTGAGTGGAGGTGTGATGGA 369
DB 102 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 121
QY 370 CTTTTCGGCTCAATATGGGACAGTGGAGAAATGGGAACAGTCAACACAGACACAGAAC 429
DB 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141
QY 430 GCGTTGTCAAGCTCATATGCAACAGAGAGAGCAAAAATAGCCATGGAGAAGCTA 489
DB 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161
QY 490 AGCGGCATCAGTTTGAGAACTACTCTTCAAGATTCTCATCATCCCGGATGACAGGTG 549
DB 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 181
QY 550 AGCTCCCTTCGCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC 609
DB 182 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201
QY 610 CACGCCCTCGGGGCACTTCTCAGGCCGACAGATTGATTTCCGCTCGGATCCTGTC 669
DB 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221
QY 670 CCCACCCAGTTGTTGGTGCCCATCATCGGAAAGAGGGCTTGACCATTAAGAACATCACT 729
```

```
DB 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAAG 789
DB 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261
QY 790 CTTGTCAACATCCATGCCACCCAGAGGGAGCTTCTGAAGCATCCCGCATGTTCTTGA 849
DB 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281
QY 850 ATCATGCAAGAGAGGAGATGACGACCAACATAGCCGAGAGATTCCTCTGAAATCTTG 909
DB 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 301
QY 910 GCACATATGCTTGGTTGGAGAGACTGATTGAAAAGAGGACGAAATTTGAAGAAATTT 969
DB 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 321
QY 970 GAACATGAACAGGACCAAGATAACAATCTCATCTTTGACAGGATTTGAGCATATACAAC 1029
DB 322 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 341
QY 1030 CCGGAAAGAACCATCACTGTGAAGGACACAGTTCAGGCCTGTGCGAGTCTGAGATAGAG 1089
DB 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361
QY 1090 ATTATGAAGAGCTGCGTGGAGGCTTTGAAAATGATATGCTGGCTGTAAAC 1140
DB 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 378

RESULT 4
US-10-313-986-501
; Sequence 501, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-501

Alignment Scores:
Pred. No.: 2.16e-314 Length: 587
Score: 356.00 Matches: 356
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 33.05% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-8 (1-3283) x US-10-313-986-501 (1-587)

QY 73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCCACCGCCGACGACCTCCGGCAG 132
DB 1 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArgGln 20
QY 133 CTCTTTGGGACAGGAGCTGCCCTCGGGGACAGGTCTCTCTGAAGTCCGGCTACGCC 192
DB 21 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyrAla 40
QY 193 TTGCTGACTATACCCCGACAGAACTCGGCGCATCCGCGCCATCGAGACCCCTCTCGGGTAAA 252
```


Db 41 PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys 60
 QY 253 GTGAAATTGATCGGGAATCATGGAAGTTGATTACTCAGTCTCTAATAAGCTAAGGAGC 312
 Db 61 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 80
 QY 313 AGGAAATTCAGATTGAAACATCCCTCCTCCTCAGTCAGTCAGGAGGTGTGATGAGTT 372
 Db 81 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu 100
 QY 373 TTGGCTCAATATCGGACAGTGGAGATGTGGAACTCAACACAGACACAGAAACCGCC 432
 Db 101 LeuAlaGlnTyrGlyThrValGluAsnValGluInValAsnThrAspThrGluThra 120
 QY 433 GTTGCTCACTCATATGCAACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
 Db 121 ValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer 140
 QY 493 GGGCATCAGTTTTCAGAACTACTCTTCAAGATTTCCTACATCCCGATCAAGAGTGAGC 552
 Db 141 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluValSer 160
 QY 553 TCCTCTTCGCTCCTCAGGAGCCAGCGTGGGACCACTCTTCCCGGAGCAGAGCCAC 612
 Db 161 SerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 180
 QY 613 GCCCTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTCGGATCTCGTCCCC 672
 Db 181 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 200
 QY 673 ACCAGTTTGTGGTCCCATCATCGGAAAGAGGGCTTGACCATAAAGAACATCACTAAG 732
 Db 201 ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys 220
 QY 733 CAGACCCAGTCCCGGCTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAGCCT 792
 Db 221 GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro 240
 QY 793 GTCAACATCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATTTGAAATC 852
 Db 241 ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle 260
 QY 853 ATCAGAAAGAGCAGATGAGACCAACTAGCCGAGAGATCTCTCTGAAATCTTGSCA 912
 Db 261 MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAla 280
 QY 913 CACAATGGCTGTGGTGGAGACTGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
 Db 281 HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGlu 300
 QY 973 CATGAACAGGGACCAAGATACAACTCTATCTTTGACGATTTTCAGCATATACAAACCG 1032
 Db 301 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 320
 QY 1033 GAAAGAACCATCATCTGTAAGGCGACAGTTGAGGGCTGTGCAGTGTGAGATAGAGATT 1092
 Db 321 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 340
 QY 1093 ATCAAGAGCTGCTGAGCGCTTGTGAATGATGATGCTGCTGTTAAC 1140
 Db 341 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 356

RESULT 5

US-09-764-864-1119
 ; Sequence 1119, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1119
 ; LENGTH: 171
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-764-864-1119

Alignment Scores:

Pred. No.: 5,13e-136 Length: 171
 Score: 160.00 Matches: 160
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.86% Indels: 0
 DE: 9 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-764-864-1119 (1-171)

QY 1258 GGCGCCATCATCGGGAAGAGGGGGCACACATCAACAGCTGGCGAGATTCCCGGAGCC 1317
 Db 12 GlyAlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
 QY 1318 TCTATCAAGATTGCCCCCTCGGAGCCGAGACGTCAGGAAAGGATGCTCATCATACC 1377
 Db 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51
 QY 1378 GGGCCACCCGGAAGCCAGTTCAAGGCCAGGACCGGATCTTTGGGAAACTGAAAGAGGAA 1437
 Db 52 GlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
 QY 1438 AACTCTTTAAACCCCAAGAAAGAGTGAAGCTGGAAGGCGCATATCAGAGTCCCTCTTCC 1497
 Db 72 AsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
 QY 1498 ACAGCTGGCGGGTGTATGCGCAAGGTGCAAGACCGTGAACGAACTGCAGAACTTAACC 1557
 Db 92 ThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
 QY 1558 AGTCGAGAAGTCTGCTGCTGTCGACCAACGCCAGATGAAAATGAGGAAAGTATCTC 1617
 Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
 QY 1618 AGAATTATCGGCACCTCTTCTGCTAGCAGACTGCACAGCGCAAGATCAGGAAATTTCTA 1677
 Db 132 ArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
 QY 1678 CAACAGTGAAGCAGCAGAGCAGAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAG 1737
 Db 152 GlnGlnValLysGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171

RESULT 6

US-09-764-864-1117
 ; Sequence 1117, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1117
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (71)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (147)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (160)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (164)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (165)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (168)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-764-864-1117

Alignment Scores:
 Pred. No.: 3.6e-54 Length: 192
 Score: 70.00 Matches: 70
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.50% Indels: 0
 DB: 9 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-764-864-1117 (1-192)

QY 781 GCAGAGACCTGTCACATCCATGCGCCAGAGGGGACTTCTCAAGCATGCCGCGATG 840
 Db 1 AlaGluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMet 20
 QY 841 ATTCTTGAATCATGACAGAAAGGCGAGATGAGCCAACTAGCCGAGAGATTCTCTG 900
 Db 21 IleLeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeu 40
 QY 901 AAATCTTGGCACATCGCTGTGGAGACTGATTGAAAGAGAGCGAGAAATTG 960
 Db 41 LysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 60
 QY 961 AAGAAATTTGACATGAACACAGGACCAAG 990
 Db 61 LysLysIleGluHisGluThrGlyThrLys 70

RESULT 7

US-09-764-864-1536
 ; Sequence 1536, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT223
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1536
 ; LENGTH: 171
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (55)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (170)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-764-864-1536

Alignment Scores:
 Pred. No.: 1.59e-50 Length: 171
 Score: 66.00 Matches: 66
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.13% Indels: 0

DB: 9 Gaps: 0

US-09-270-437D-8 (1-3283) x US-09-764-864-1536 (1-171)

QY 1480 ATCAGAGTGCCTCTTCCACAGCTGGCGGTGATTGGCAAAGGTGGCAAGACCGTGAAC 1539
 Db 86 IleArgValProSerSerThrAlaGlyArgValIleGlyLysGlyLysThrValAsn 105
 QY 1540 GAACCTGCAGAACTTAACAGTGCAGAACTCATCTGCTGTCACCAAAAGCCAGATGAA 1599
 Db 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125
 QY 1600 AATCAGAAAGTATGTCGTGAGAAATATATCGGCACACTCTTTCTAGCCAGACTGCACAGCGC 1659
 Db 126 AsnGluGluValIleValArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArg 145
 QY 1660 AAGATCAGGGAATGTGA 1677
 Db 146 LysIleArgGluIleVal 151

RESULT 8

US-10-262-445-40
 ; Sequence 40, Application US/10262445
 ; Publication No. US20040014058A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alsobrook II, John
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Catterton, Elina
 ; APPLICANT: Chant, John
 ; APPLICANT: Chaudhuri, Amitabha
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Mezes, Peter
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Ooi, Chean Eng
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Rieger, Daniel
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Taupier Jr., Raymond J.
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Zhong, Hailong
 ; APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
 FILE REFERENCE: 21402-462D
 CURRENT APPLICATION NUMBER: US/10/262,445
 CURRENT FILING DATE: 2002-10-01
 PRIOR APPLICATION NUMBER: 60/327,454
 PRIOR FILING DATE: 2001-10-05
 PRIOR APPLICATION NUMBER: 60/327,917
 PRIOR FILING DATE: 2001-10-09
 PRIOR APPLICATION NUMBER: 60/328,029
 PRIOR FILING DATE: 2001-10-09
 PRIOR APPLICATION NUMBER: 60/328,056
 PRIOR FILING DATE: 2001-10-09
 PRIOR APPLICATION NUMBER: 60/328,849
 PRIOR FILING DATE: 2001-10-12
 PRIOR APPLICATION NUMBER: 60/329,414
 PRIOR FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: 60/330,142
 PRIOR FILING DATE: 2001-10-17
 PRIOR APPLICATION NUMBER: 60/341,058
 PRIOR FILING DATE: 2001-10-22
 PRIOR APPLICATION NUMBER: 60/343,629
 PRIOR FILING DATE: 2001-10-24
 PRIOR APPLICATION NUMBER: 60/349,575
 PRIOR FILING DATE: 2001-10-29
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 133

```
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-445-40
Alignment Scores:
Pred. No.: 1.35e-30 Length: 555
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.09% Indels: 0
DB: 15 Gaps: 0
US-09-270-437D-8 (1-3283) x US-10-262-445-40 (1-555)
Qy 325 ATTGAAACATCCCTCTCCTACCTGCGAGTGTGGATGCGACTTTGGCTCAATAT 384
Db 86 IleArgAsnIleProHisLeuGlnTrpGluValLeuAspGlyLeuAlaGlnTyr 105
Qy 385 GGCACAGTGGAGATGTGGAAACAAGTCAACACACACACAGAAACCGCGTGTCAACGTC 444
Db 106 GlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValValAsnVal 125
Qy 445 ACATATGCAACA 456
Db 126 ThrTyrAlaThr 129
RESULT 9
US-10-313-986-499
; Sequence 499, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCES: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 499
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-499
Alignment Scores:
Pred. No.: 1.54e-08 Length: 20
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.86% Indels: 0
DB: 15 Gaps: 0
US-09-270-437D-8 (1-3283) x US-10-313-986-499 (1-20)
Qy 1324 AAGATTGCCCTCGGAAGGCCAGACGTCAGCGAAGGATGGTCATCATCACCGGGCCA 1383
Db 1 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 20
RESULT 10
US-09-764-864-1534
; Sequence 1534, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1534
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (24)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1534
Alignment Scores:
Pred. No.: 1.35e-08 Length: 47
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.86% Indels: 0
DB: 9 Gaps: 0
US-09-270-437D-8 (1-3283) x US-09-764-864-1534 (1-47)
Qy 643 ATTGATTCCCGCTCGGATCCCTGGTCCGCCACCCAGTTTGTGGCCATCATCGGAAG 702
Db 1 IleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIleIleGlyLys 20
RESULT 11
US-09-873-637-20
; Sequence 20, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-20
Alignment Scores:
Pred. No.: 1.09e-07 Length: 48
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 9 Gaps: 0
US-09-270-437D-8 (1-3283) x US-09-873-637-20 (1-48)
Qy 1501 GCTGGCGGGTGATTCGCAAGGTGGCAAGACCGTGAACGAACTGCAGACTTAACC 1557
Db 8 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 26
RESULT 12
US-10-117-982-476
; Sequence 476, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Ranger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
```

US 9,304,132; Sequence 1532, Application US/09764864

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Search completed: August 6, 2004, 14:13:30
Job time : 170.275 secs

us-09-270-437d-8.oligo806.rapb

Fri Aug 6 14:50:36 2004

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 49.9137 Seconds
(without alignments)
12653.723 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcagcgaggaggcgagga.....aaccttgaaatgtttattt 3283

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=klp
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat_06082004_141221_1829/app_query.fasta_1.5582
-DB=PIR 78 -QMFT=fastan -SUFFIX=oligo806.rpr -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437@cgn_1_134@runat_06082004_141221_1829 -NCPH=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR 78:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	0.8	319	2 F75420	hypothetical prote
C 2	9	0.8	366	2 AB1422	probable GRP-bindi
C 3	9	0.8	366	2 AH1796	probable GRP-bindi
C 4	9	0.8	398	1 WZBEN3	protein-serine/thr
C 5	9	0.8	431	2 T29850	hypothetical prote
C 6	9	0.8	491	1 FGHUB	fibrinogen beta ch
C 7	9	0.8	696	2 A29635	transcription fact
C 8	9	0.8	788	2 JS0747	regulatory protein
C 9	9	0.8	1226	2 T49915	pre-mRNA splicing
C 10	9	0.8	1234	2 S52099	phospholipase C be
C 11	9	0.8	2338	2 I73957	kinase-related pro
C 12	9	0.8	2347	1 TVHURS	kinase-related pro
C 13	8	0.8	63	2 T12121	NADH dehydrogenase
C 14	8	0.8	121	2 F72580	hypothetical prote

C 15	8	0.8	125	2 A71245	hypothetical prote
C 16	8	0.7	137	2 S37353	nodulin (clone GmE
C 17	8	0.8	137	2 G75471	hypothetical prote
C 18	8	0.7	146	2 T16341	hypothetical prote
C 19	8	0.7	150	2 T17206	hypothetical prote
C 20	8	0.8	157	2 A31293	shikimate kinase h
C 21	8	0.7	161	1 S76604	hypothetical prote
C 22	8	0.7	171	2 B89975	conserved hypothet
C 23	8	0.8	183	2 AF1883	hypothetical prote
C 24	8	0.7	202	2 T46586	ribosomal protein
C 25	8	0.7	205	2 T34724	probable membrane
C 26	8	0.8	208	2 T16953	hypothetical prote
C 27	8	0.8	219	2 A99194	iron (III) ABC tra
C 28	8	0.7	220	2 A36288	proline-rich prote
C 29	8	0.7	242	2 G96994	glycerol uptake fa
C 30	8	0.8	257	2 C96965	transcription regu
C 31	8	0.7	267	2 D83133	probable permease
C 32	8	0.7	279	2 D82281	ferric vibriobacti
C 33	8	0.7	285	2 H70781	hypothetical prote
C 34	8	0.7	301	2 G93182	hypothetical prote
C 35	8	0.7	311	2 A38558	interferon respons
C 36	8	0.7	315	2 A84634	hypothetical prote
C 37	8	0.8	316	2 E71812	transaldolase - He
C 38	8	0.8	320	2 AF1892	[Nife] uptake hydr
C 39	8	0.7	324	2 G90896	probable transcrip
C 40	8	0.7	326	2 S56534	hypothetical 36.9K
C 41	8	0.7	326	2 D31287	hypothetical prote
C 42	8	0.7	326	2 G86128	hypothetical prote
C 43	8	0.7	327	2 H85720	hypothetical prote
C 44	8	0.7	334	2 T36485	probable araC-famI
C 45	8	0.7	343	2 C97656	periplasmic iron-b

ALIGNMENTS

RESULT 1
F75420
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
R:Accession: F75420
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75420
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <WHI>
A:Cross-references: GB:AE001971; GB:AE000513; NID:G6458972; PID:AAF10810.1; PID:G64589
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1229
A:Map position: 1

Alignment Scores:				
Pred. No.:	11.5	Length:	319	
Score:	9.00	Matches:	9	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	0.85%	Indels:	0	
DB:	2	Gaps:	0	
US-09-270-437D-8 (1-3283) x F75420 (1-319)				
QY	1079	GCATCGGCACAGCCCTCACTGTGCC	1053	
Db	7	AlaLeuAlaGlnAlaSerThrValPro	15	
RESULT 2				
AB1422				

probable GTP-binding protein homolog lmo2779 [imported] - Listeria monocytogenes (strain C) Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AB1422
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1422
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-366 <GLA>
A;Cross-references: GB:NC 003210; PIDN:CAD00992.1; PID:g16412279; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c
A;Gene: lmo2779

Alignment Scores:
Pred. No.: 11.3 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.84% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x AB1422 (1-366)

QY 1099 AAGTCGCTGAGGCGCTTTGAAATGAT 1125

Db 167 LysLeuArgGluAlaPheGluAsnAsp 175

RESULT 3

AH1796

Probable GTP-binding protein homolog lin2919 [imported] - Listeria innocua (strain Clip1 C) Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1796
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-366 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC98144.1; PID:g16415460; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
C;Gene: lin2919
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c

Alignment Scores:
Pred. No.: 11.3 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.84% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x AH1796 (1-366)

QY 1099 AAGTCGCTGAGGCGCTTTGAAATGAT 1125

Db 167 LysLeuArgGluAlaPheGluAsnAsp 175

RESULT 4

WBEN3

Protein-serine/threonine kinase (EC 2.7.1.1) - suid herpesvirus 1 (strain NIA-3)
N;Alternate names: UL13 protein
C;Species: suid herpesvirus 1
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 11-Jun-1999
C;Accession: B42744
R;de Wind, N.; Domen, J.; Berns, A.
J. Virol. 66, 5200-5209, 1992
A;Title: Herpesviruses encode an unusual protein-serine/threonine kinase which is nonessential for replication.
A;Reference number: A42744; MUID:92365105; PMID:1323689
A;Accession: B42744
A;Molecule type: DNA
A;Residues: 1-398 <DEW>
A;Cross-references: GB:M94870; NID:g334092; PIDN:AAA47481.1; PID:g334094
C;Genetics:
C;Gene: UL13
C;Superfamily: herpesvirus protein-serine/threonine kinase; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;78-378/Domain: protein kinase homology <KIN>
F;86-93/Region: protein kinase ATP-binding motif
F;103/Active site: Lys #status predicted

Alignment Scores:
Pred. No.: 11.2 Length: 398
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.84% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x WBEN3 (1-398)

QY 2 GCAGCGAGGAGCGGAGGCGCGGG 28

Db 119 AlaAlaGluGluAlaArgSerAlaGly 127

RESULT 5

T29850

Hypothetical protein C49C8.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29850
R;Johnson, D.; Bradshaw, H.
submitted to the EMBL Data Library, June 1996
A;Description: The sequence of C. elegans cosmid C49C8.
A;Reference number: Z20696
A;Accession: T29850
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431 <JOH>
A;Cross-references: EMBL:U61945; PIDN:AAB03126.1; GSPDB:GN00022; CESP:C49C8.5
A;Experimental source: strain Bristol N2; clone C49C8
C;Genetics:
A;Gene: CESP:C49C8.5
A;Map position: 4
A;Introns: 17/3; 66/2; 106/3; 151/2; 187/1; 233/2; 302/3; 329/1; 351/2; 377/2

Alignment Scores:
Pred. No.: 11 Length: 431
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x T29850 (1-431)

QY 2461 AAGGAGCGGTATCCCTTCTGTCACA 2435

Db 3 LysGluThrValSerLeuLeuLeuThr 11

RESULT 6

FGHUB
fibrinogen beta chain precursor [validated] - human
N;Alternate names: coagulation factor I
N;Contents: fibrinopeptide B
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 31-Mar-1993 #text_change 08-Dec-2000
C;Accession: B43568; A90469; B90469; I37389; A94433; A90437; A94309; G54223; A03121; B37
R;Chung, D.W.; Harris, J.E.; Davie, E.W.
Adv. Exp. Med. Biol. 281, 39-48, 1990
A;Title: Nucleotide sequences of the three genes coding for human fibrinogen.
A;Reference number: A43568; MUID:91344740; PMID:2102623
A;Accession: B43568
A;Molecule type: DNA
A;Residues: 9-191, 'P', 193-491 <CHU>
R;Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.
Biochemistry 22, 3244-3250, 1983
A;Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonu
A;Reference number: A90469; MUID:83283433; PMID:6688356
A;Accession: A90469
A;Molecule type: DNA
A;Residues: 1-38 <CH1>
A;Accession: B90469
A;Molecule type: mRNA
A;Residues: 9-191, 'A', 193-491 <CH2>
A;Cross-references: GB:J00129; NID:gi182429; PIDN:AA52429.1; PID:gl82430
R;Huber, P.; Dalmon, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.
Nucleic Acids Res. 15, 1615-1625, 1987
A;Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.
A;Reference number: I37389; MUID:87146483; PMID:3029722
A;Accession: I37389
A;Status: translated from GB/EMBL/DDDBJ
A;Molecule type: DNA
A;Residues: 1-38 <HUB>
A;Cross-references: EMBL:X05018; NID:G31400; PIDN:CAA28674.1; PID:g31401
R;Henschen, A.; Lottspeich, F.; Southan, C.; Topfer-Petersen, E.
in Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe
A;Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v
A;Reference number: A94433
A;Contents: carbohydrate binding
A;Accession: A94433
A;Molecule type: protein
A;Residues: 31-137, 'QS', 140-144, 'QF', 147-491 <HEN>
R;Watt, K.W.K.; Takagi, T.; Doolittle, R.F.
Biochemistry 18, 68-76, 1979
A;Title: Amino acid sequence of the beta chain of human fibrinogen.
A;Reference number: A90437; MUID:79124640; PMID:420779
A;Accession: A90437
A;Molecule type: protein
A;Residues: 31-144, 'QF', 147-231, 'D', 233-330, 'E', 332-491 <WAT>
R;Blomback, B.; Hessel, B.; Hogg, D.
Thromb. Res. 8, 639-658, 1976
A;Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.
A;Reference number: A94309; MUID:76225080; PMID:936108
A;Contents: disulfide bonds
A;Accession: A94309
A;Molecule type: protein
A;Residues: 31-112, 'E', 114-137, 'QS', 140-144, 'QF', 147-148 <BLO>
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
Biochemistry 33, 1988-1993, 1994
A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipopr
A;Reference number: A54223; MUID:94162201; PMID:8117655
A;Accession: G54223
A;Molecule type: protein
A;Residues: 164-174 <KUN>
A;Note: identification of tryptic peptides from high-density lipoproteins
R;Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.
Ann. N. Y. Acad. Sci. 408, 28-43, 1983
A;Title: Covalent structure of fibrinogen.
A;Reference number: A90037; MUID:83254370; PMID:6575689
A;Contents: annotation; review, disulfide bonds
R;Gardlund, B.; Hessel, B.; Marguerie, G.; Murano, G.; Blomback, B.
Eur. J. Biochem. 77, 595-610, 1977

A;Title: Primary structure of human fibrinogen. Characterization of disulfide-containing
A;Reference number: A91249; MUID:77245999; PMID:891553
R;Doolittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.; (C
in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ottessen, M., Folt
A;Title: The structures of fibrinogen and fibrin.
A;Reference number: A94437
A;Contents: annotation; disulfide bonds
R;Doolittle, R.F.
Annu. Rev. Biochem. 53, 195-229, 1984
A;Title: Fibrinogen and fibrin.
A;Reference number: A90041; MUID:84305751; PMID:6383194
A;Contents: annotation; review, EM structure, polymerization, ligands
R;Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.
Ann. N. Y. Acad. Sci. 408, 449-456, 1983
A;Title: Cloning of fibrinogen genes and their cDNA.
A;Reference number: A90038; MUID:83254384; PMID:6575700
A;Contents: annotation
R;Kirschbaum, N.E.; Budzynski, A.Z.
J. Biol. Chem. 265, 13669-13676, 1990
A;Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-t
A;Reference number: A37117; MUID:90337977; PMID:2143188
A;Contents: annotation; hementin cleavage site
A;Note: hementin, a protease from Haemateria ghilianii, the giant South American leech,
C;Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cleav
ization sites responsible for the formation of the soft clot.
C;Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabil
ger) and between alpha chains (weaker) of different monomers.
C;Comment: All fibrinogen chains are synthesized in the liver.
C;Genetics:
A;Gene: GDB:FG8
A;Cross-references: GDB:119130; OMIM:134830
A;Map position: 4q28-q28
A;Intons: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2
A;Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:FI
ins are contained in the core. Two three-chain coiled coils emerge from this core and c
from the distal domain nodes.
C;Function:
A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into
A;Pathway: blood coagulation
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulf
C;Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; Pyroglutamic a
F1-30/Domain: (or 4-30 or 15-30) signal sequence #status experimental <SIG>
F31-491/Product: fibrinogen beta chain #status experimental <MAT>
F31-44/Product: fibrinopeptide B #status experimental <APT>
F45-491/Product: fibrin beta chain #status experimental <FGB>
F45-47/Region: polymerization site
F59-228/Domain: fibrinogen disulfide ring homology <FDR>
F238-487/Domain: fibrinogen beta/gamma homology <FBG>
F31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime
F44-45/cleavage site: Arg-Gly (thrombin) #status experimental
F95/Disulfide bonds: interchain (to alpha-55) #status experimental
F106/Disulfide bonds: interchain (to alpha-68) #status experimental
F110/Disulfide bonds: interchain (to gamma-45) #status experimental
F223/Disulfide bonds: interchain (to alpha-164) #status experimental
F227/Disulfide bonds: interchain (to gamma-161) #status experimental
F231-316, 241-270, 424-437/Disulfide bonds: #status experimental
F394/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:
Pred. No.: 10.8 Length: 491
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x FGHUB (1-491)

QY 3042 CTCTCTCTGTGTGTGTGTCTCTGTAA 3016
|||||
Db 21 LeuLeuLeuCysValPheLeuValIys 29

RESULT 7
A29635
transcription factor Sp1 - human (fragment)
N:Alternate names: finger protein ZNF76
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 05-Nov-1999
C:Accession: A29635; G44256
R:Kadonaga, J.T.; Carner, K.R.; Masiaz, F.R.; Tjian, R.
Cell 51, 1079-1090, 1987
A:Title: Isolation of cDNA encoding transcription factor Sp1 and functional analysis of
A:Reference number: A29635; MUID:98080466; PMID:3319186
A:Accession: A29635
A:Molecule type: mRNA
A:Residues: 1-696 <KAD>
A:Cross-references: GB:J01133; NID:g339517; PID:AAA61154.1; PID:g339518
R:Ragoussis, J.; Senger, G.; Mockridge, I.; Sanseau, P.; Ruddy, K.; Sheer, D.
Genomics 14, 673-679, 1992
A:Title: A testis-expressed Zn finger gene (ZNF76) in human 6p21.3 centromeric to the M1
A:Reference number: A44256; MUID:93052398; PMID:1427894
A:Accession: G44256
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 569-598 <RAG>
A:Experimental source: T-cell line CEM
A:Note: sequence extracted from NCBI backbone (NCBIP:125980)
C:Genetics:
A:Gene: GDB:SP1
A:Cross-references: GDB:127453; OMIM:189906
A:Map position: 19q13.1-19q13.3
C:Keywords: DNA binding; transcription regulation; zinc finger

Alignment Scores:
Pred. No.: 10.3 Length: 696
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0
Gaps: 0
DB:

US-09-270-437D-8 (1-3283) x A29635 (1-696)

QY 893 ATCTCTCGGCTAGTTGGTCTCATCT 867
|||||
DB 216 lleserSerAlaSerLeuValSerSer 224

RESULT 8
JS0747
regulatory protein Sp1 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C:Accession: JS0747; S25287
R:Inataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami,
submitted to JIPID, September 1992
A:Reference number: JS0747
A:Accession: JS0747
A:Molecule type: mRNA
A:Residues: 1-788 <INA>
A:Cross-references: DBJ:D12768; NID:g220911; PID:BAA02235.1; PID:d1002730; PID:g220912
R:Inataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami,
EMBO J. 11, 3663-3671, 1992
A:Title: Two regulatory proteins that bind to the basic transcription element (BTE), a C
A:Reference number: S25287; MUID:93010958; PMID:11356762
A:Accession: S25287
A:Molecule type: mRNA
A:Residues: 1-122, 'L', 124-311, 'A', 313-788 <IM2>
C:Keywords: DNA binding; transcription regulation

Alignment Scores:
Pred. No.: 10.1 Length: 788
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0

DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x JS0747 (1-788)

QY 893 ATCTCTCGGCTAGTTGGTCTCATCT 867
|||||
DB 308 lleserSerAlaSerLeuValSerSer 316

RESULT 9
T49915
pre-mRNA splicing factor ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
N:Alternate names: protein T24H18.180
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49915
R:Bevan, M.; Robben, J.; Grymonprez, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Ruddy,
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25024
A:Accession: T49915
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1226 <BEV>
A:Cross-references: EMBL:AL353013; GSPDB:GN00063; ATSP:T24H18.180
A:Experimental source: cultivar Columbia; BAC clone T24H18
C:Genetics:
A:Gene: ATSP:T24H18.180
A:Map position: 5
A:Introns: 2/3; 40/1; 111/1; 123/3; 183/3; 231/1; 279/3; 313/2; 349/3; 409/1; 485/3; 544/3

Alignment Scores:
Pred. No.: 9.52 Length: 1226
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0
Gaps: 0
DB:

US-09-270-437D-8 (1-3283) x T49915 (1-1226)

QY 499 GATGCCGCTAGTCTCCATGGCTA 473
|||||
DB 216 AspAlaArgLeuAlaSerProTirpueu 224

RESULT 10
S52099
phospholipase C beta 3 - human
N:Alternate names: phospholipase c beta-3, phosphoinositide-specific
C:Species: Homo sapiens (man)
C>Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999
C:Accession: S52099; A56833; A56854
R:Lagarcrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld
submitted to the EMBL Data Library, September 1994
A:Description: Genomic organization and complete cDNA sequence of the human phosphoinosi
A:Reference number: S52099
A:Accession: S52099
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1234 <LAG>
A:Cross-references: EMBL:Z37573
R:Lagarcrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld
Genomics 26, 467-472, 1995
A:Title: Genomic organization and complete cDNA sequence of the human phosphoinositide-s
A:Reference number: A56833; MUID:93331781; PMID:7607669
A:Accession: A56833
A:Molecule type: DNA
A:Residues: 1-200 <LA2>
A:Cross-references: GB:Z37544
R:Sink, R.J.; Geurts van Kessel, A.
Genomics 25, 568-569, 1995
A:Title: Localization of the human phosphatidylinositol-specific phospholipase C beta-3
A:Reference number: A56854; MUID:93309527; PMID:7789993
A:Accession: A56854
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA
A:Residues: 1-193 <SIN>

C:Gene: GDB:PLCB3
A:Cross-references: GDB:386061; OMIM:600230

A:Map position: 11q13-11q13

A:Introns: 33/3; 59/3; 82/3; 129/3; 156/2; 174/2; 199/3; 233/2; 289/3; 338/1; 418/2; 446/3; 1119/2; 1138/3; 1167/3

C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase I; 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom

F:319-468/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
F:589-709/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Alignment Scores:

Pred. No.:	9.51	1234
Score:	9.00	9
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	0.85%	Indels: 0
DB:	2	Gaps: 0

US-09-270-437D-8 (1-3283) x S52099 (1-1234)

QY 217 AGTCTGTCGGGTGAGTCCACGAGG 191

DB 1099 SerSerGlyArgGlySerProArgArg 1107

RESULT 11

I73957

Kinase-related protein c-ros-1 precursor - rat

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 02-Jun-2000

A:Accession: I73957; I56752; I73956

R:Matsumura, H.; Shibuya, M.

J. Virol. 64, 2117-2125, 1990

A:Title: Tissue-specific expression of rat c-ros-1 gene and partial structural similarity

A:Reference number: I56752; MUID:90219211; PMID:2139140

A:Accession: I73957

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2338 <RES>

A:Cross-references: GB:M35106; NID:g203599; PIDN:AAA40968.1; PID:g203600

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-430,452-2338 <RE2>

A:Cross-references: GB:M35104; NID:g203595; PIDN:AAA40966.1; PID:g203596

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-430,452-1872, 'AC', 1875 <RE3>

A:Cross-references: GB:M35105; NID:g203597; PIDN:AAA40967.1; PID:g203598

C:Superfamily: kinase-related protein ros; LDL receptor WYTD-containing repeat homology; C:KeyWords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related

ific protein kinase

F:753-793/Domain: LDL receptor WYTD-containing repeat homology <YW3>

F:1935-2214/Domain: protein kinase homology <KIN>

F:1943-1951/Region: protein kinase ATP-binding motif

Alignment Scores:

Pred. No.:	8.69	2338
Score:	9.00 <td>9</td>	9
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	0.84%	Indels: 0
DB:	2	Gaps: 0

US-09-270-437D-8 (1-3283) x I73957 (1-2338)

QY 569 AGCGAGCCAGCGTGGGACCACTCTT 595

DB 656 SerGlnProSerValGlyThrIleu 664

RESULT 12

TVHURS

Kinase-related protein ros-1 precursor - human

N:Alternate names: protein-tyrosine kinase mcf3 (activated ros-1)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1998 #sequence revision 07-Oct-1994 #text_change 11-Jun-1999

R:Accession: A35512; A25223; A24421; A33081

R:BiSchmeier, C.; O'Neill, K.; Riggs, M.; Wigler, M.

Proc. Natl. Acad. Sci. U.S.A. 87, 4799-4803, 1990

A:Title: Characterization of ROS1 cDNA from a human glioblastoma cell line.

A:Reference number: A35512; MUID:90280463; PMID:2352949

A:Accession: A35512

A:Molecule type: mRNA

A:Residues: 1-2212, 'N', 2214-2227, 'QC', 2229-2347 <BIR>

A:Cross-references: GB:M34353

A:Experimental source: glioblastoma cell line SW-1088

R:Matsumura, H.; Wang, L.H.; Shibuya, M.

Mol. Cell. Biol. 6, 3000-3004, 1986

A:Title: Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma virus encoded

A:Reference number: A25223; MUID:87064611; PMID:3023556

A:Accession: A25223

A:Molecule type: DNA

A:Residues: 1790-2245, 'KFDSEFSFSRCTVN' <MA2>

A:Cross-references: GB:M13368

A:Experimental source: Placenta

A:Note: the differences after residue 2245 result from the authors' misinterpretation o

R:BiSchmeier, C.; Birnbaum, D.; Waitches, G.; Fasano, O.; Wigler, M.

Mol. Cell. Biol. 6, 3109-3116, 1986

A:Title: Characterization of an activated human ros gene.

A:Reference number: A24421; MUID:87064625; PMID:3785223

A:Accession: A24421

A:Molecule type: mRNA

A:Residues: 1854-2261, 'A', 2263-2347 <BI2>

A:Cross-references: GB:M13880; NID:g337482; PIDN:AAA36580.1; PID:g337483

A:Experimental source: tumor cells

A:Note: the mcf3 oncogene was formed by DNA rearrangement involving fusion of at least

C:Genetics:

A:Gene: GDB:ROS1

A:Cross-references: GDB:I20351; OMIM:165020

A:Map position: 6q22-q22

A:Introns: 185/1; 188/1; 1926/2; 1980/3; 2002/2; 2045/3; 2078/2; 2145/2; 2190/2

C:Superfamily: kinase-related protein ros; LDL receptor WYTD-containing repeat homology;

C:KeyWords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;

F:1-36/Domain: signal sequence #status predicted <SIG>

F:37-2347/Product: kinase-related protein ROS1 #status predicted <MAT>

F:37-1859/Domain: extracellular #status predicted <EXT>

F:335-378/Domain: LDL receptor WYTD-containing repeat homology <YML>

F:466-503/Domain: LDL receptor WYTD-containing repeat homology <YWA>

F:715-757/Domain: LDL receptor WYTD-containing repeat homology <YW2>

F:758-798/Domain: LDL receptor WYTD-containing repeat homology <YW3>

F:799-838/Domain: LDL receptor WYTD-containing repeat homology <YW4>

F:843-888/Domain: LDL receptor WYTD-containing repeat homology <YW5>

F:893-933/Domain: LDL receptor WYTD-containing repeat homology <YW6>

F:1532-1574/Domain: LDL receptor WYTD-containing repeat homology <YW7>

F:1860-1883/Domain: transmembrane #status predicted <TMN>

F:1884-2347/Domain: intracellular #status predicted <INT>

F:1943-2222/Domain: protein kinase homology <KIN>

F:1951-1959/Region: protein kinase ATP-binding motif

F:52,114,123,324,352,471,607,628,706,714,732,939,961,1015,1087,1090,1211,1272,1330,1458,

F:1980/Active site: Lys #status predicted

F:2110,2114,2115/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #sta

Alignment Scores:	8.69	Length:	2347
Pred. No.:	9.00	Matches:	9
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	0.84%	Gaps:	0
DB:	1		

US-09-270-437D-8 (1-3283) x TVHURS (1-2347)

```

QY 569 AGCGAGCCCGCGTGGGACCACTCTT 595
Db 661 SerGluProSerValGlyThrThrLeu 669

RESULT 13
T12121
NADH dehydrogenase 4 - Atlantic horseshoe crab mitochondrion (fragment)
C:Species: mitochondrion Limulus polyphemus (Atlantic horseshoe crab)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T12121
R:Stacion, J.L.; Daehler, L.L.; Brown, W.M.
Mol. Biol. Evol. 14, 867-874, 1997
A:Title: Mitochondrial gene arrangement of the horseshoe crab Limulus polyphemus L.: Cor
A:Reference number: Z17427; MUID:97398711; PMID:9254925
A:Accession: T12121
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-63 <STA>
A:Cross-references: EMBL:AF002647; NID:92316048; PID:92316055; PIDN:AAC47689.1
C:Genetics:
A:Genome: mitochondrion
A:Note: ND4
C:Keywords: mitochondrion

Alignment Scores:
Pred. No.: 142 Length: 63
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x T12121 (1-63)

QY 1704 TTCTGCTCTGCTGCTTCACCTG 1681
Db 39 PheLeuLeuLeuLeuHisLeu 46

RESULT 14
F72580
hypothetical protein APE1925 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72580
R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Tanaka, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72580
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <KAW>
A:Cross-references: DDBJ:AF000062; NID:95105244; PIDN:BAA80931.1; PID:di044717; PID:9510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1925

Alignment Scores:
Pred. No.: 130 Length: 121
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x F72580 (1-121)

QY 1929 CTTCCCGCCCTCTCTCGGCCCT 1906
Db 95 LeuProArgProSerSerAlaPro 102
```

```

RESULT 15
A71245
hypothetical protein PH0219 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: A71245
R:Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71245
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-125 <KAW>
A:Cross-references: GB:AP000001; NID:93236128; PIDN:BAA29288.1; PID:93256605
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0219

Alignment Scores:
Pred. No.: 129 Length: 125
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-8 (1-3283) x A71245 (1-125)

QY 1929 CTTCCCGCCCTCTCTCGGCCCT 1906
Db 28 LeuProArgProSerSerAlaPro 35

Search completed: August 6, 2004, 13:33:48
Job time : 70.9137 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 28.8809 Seconds
(without alignments)
11838.037 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaatgtttttt 3283

Scoring table:
OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 segs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_P/US09270437/runat_06082004_141220_1803/app_query.fasta_1.5582
-DB=SwissProt 42 -QPMT=fastan -SUFFIX=oligo806.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09270437 @CNC 1 1 57 @runat_06082004_141220_1803 -NCPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELEX=7

Database : SwissProt_42.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.8	398	1 KR2_PRVN3	P30662 pseudorabie
2	9	0.8	491	1 FIBB_HUMAN	P02875 homo sapien
3	9	0.8	781	1 SP1_MOUSE	O89090 mus musculu
4	9	0.8	785	1 SP1_HUMAN	P08047 homo sapien
5	9	0.8	788	1 SP1_RAT	Q01714 rattus norv
6	9	0.8	2347	1 KR05_HUMAN	P08922 homo sapien
7	8	0.8	111	1 FTSB_RALSO	O8y0b4 raltostonia s
8	8	0.7	137	1 N551_SOYBN	Q05544 glycine max
9	8	0.7	150	1 SH1B_PIG	P79339 sus scrofa
10	8	0.7	161	1 SH1B_CANFA	P79250 canis famli
11	8	0.7	171	1 Y175_STAAM	Q53719 staphylococ
12	8	0.7	202	1 RS7_NEUCR	O43105 neurospora
13	8	0.8	208	1 VXS2_CAEEL	Q10021 caenorhabdi
14	8	0.8	224	1 RS2_METWA	O8p441 methanosarc
15	8	0.8	225	1 RS2_METAC	O8tt39 methanosarc
16	8	0.7	244	1 CSM1_CHLTE	O68988 chlorobium
17	8	0.7	285	1 Y891_MYCTU	Q10551 mycobacteri
18	8	0.7	285	1 Y915_MYCBO	P59970 mycobacteri

C 19	8	0.8	303	1 GDFP_MOUSE	Q9z0j7 mus musculu
C 20	8	0.8	303	1 GDFP_RAT	Q9z0j6 rattus norv
C 21	8	0.7	311	1 RBFI_MOUSE	P22560 mus musculu
C 22	8	0.8	316	1 TAL_HELPJ	Q9zic5 helicobacte
C 23	8	0.7	326	1 YJHS_ECOLI	P39370 escherichia
C 24	8	0.8	379	1 CYB_PENFU	Q9b1x4 penthalagus
C 25	8	0.8	381	1 CYB_NORTY	O03478 notoryctes
C 26	8	0.7	386	1 SH1B_CRIGR	P46636 cricetus
C 27	8	0.7	386	1 SH1B_MOUSE	P28334 mus musculu
C 28	8	0.7	386	1 SH1B_RAT	P28564 rattus norv
C 29	8	0.7	386	1 SH1B_SPAAH	P56496 spalax leuc
C 30	8	0.7	388	1 SH1B_DIDMA	P35404 didelphis m
C 31	8	0.7	389	1 SH1B_CAVPO	O08832 cavia porce
C 32	8	0.7	390	1 SH1B_HUMAN	P28222 homo sapien
C 33	8	0.7	390	1 SH1B_PANTR	P60020 pan troglod
C 34	8	0.7	390	1 SH1B_RABIT	P49144 oryctolagus
C 35	8	0.7	395	1 UMPI_ARATH	O91k45 arabidopsis
C 36	8	0.8	407	1 IE68_HSVSA	Q01042 herpesvirus
C 37	8	0.8	411	1 IHH_MOUSE	P97812 mus musculu
C 38	8	0.7	428	1 FXB2_MOUSE	O64733 mus musculu
C 39	8	0.8	480	1 PRTP_HUMAN	P10619 homo sapien
C 40	8	0.8	498	1 VE2_HPV08	P06422 human papil
C 41	8	0.8	499	1 PITB_ECOLI	P43676 escherichia
C 42	8	0.7	500	1 CP46_HUMAN	Q9y6a2 homo sapien
C 43	8	0.7	500	1 CP46_MOUSE	Q9wvk8 mus musculu
C 44	8	0.8	504	1 ATIN_HSVBP	P30020 bovine herp
C 45	8	0.8	511	1 ACH5_CAEEL	Q23022 caenorhabdi

ALIGNMENTS

RESULT 1
KR2_PRVN3
ID KR2_PRVN3 STANDARD; PRT; 398 AA.
AC P30662;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Serine/threonine-protein kinase 2 (EC 2.7.1.-).
GN UL13 OR ULPK.
OS Pseudorabies virus (strain NIA-3) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10349;
RN [1]_TaxID=10349;
RP SEQUENCE FROM N.A.
RX MEDLINE=92365105; PubMed=1323689;
RA de Wind N., Domen J., Berns A.;
RT "Herpesviruses encode an unusual protein-serine/threonine kinase
RT which is nonessential for growth in cultured cells.";
RL J. Virol. 66:5200-5209(1992).
CC -!- FUNCTION: PROTEIN-SERINE/THREONINE KINASE NONESSENTIAL FOR GROWTH
CC IN CULTURED CELLS.
CC -!- PTM: Autophosphorylated (possible).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M94870; AAA47481.1; -;
DR PIR: B42744; WZBN3.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR008271; Ser_thr_kinase.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Manganese.
KW POLY-GLY.
FT DOMAIN 4 9

FT DOMAIN 80 398 PROTEIN KINASE.
 FT NP BIND 86 94 ATP (BY SIMILARITY).
 FT BINDING 103 103 ATP (BY SIMILARITY).
 FT ACT SITE 194 194 BY SIMILARITY.
 SQ SEQUENCE 398 AA; 41416 MW; DSC69AD75E42309B CRC64;

Alignment Scores:
 Pred. No.: 4.89 Length: 398
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.84% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x KR2_PRVN3 (1-398)

QY 2 GCAGCGGAGGCGAGGAGCGCGCGG 28

Db 119 AlaAlaGluGluAlaArgSerAlaGly 127

RESULT 2

FIBB HUMAN
 ID FIBB HUMAN STANDARD; PRT; 491 AA.
 AC P02675; 9.00
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].
 GN FGB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91344740; PubMed=2102623;
 RA Chung D.W., Harris J.E., Davie E.W.;
 RT "Nucleotide sequences of the three genes coding for human
 fibrinogen.";
 RL Adv. Exp. Med. Biol. 281:39-48(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83283433; PubMed=6688355;
 RA Chung D.W., Que B.G., Rixon M.W., Mace M. Jr., Davie E.W.;
 RT "Characterization of complementary deoxyribonucleic acid and genomic
 deoxyribonucleic acid for the beta chain of human fibrinogen.";
 RL Biochemistry 22:3244-3250(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Chung D.W., Harris J.E., Davie E.W.;
 RT "Nucleotide sequences of the three genes coding for human
 fibrinogen.";
 RL (in) Liu C.Y., Chien S. (eds.);
 RL Fibrinogen, thrombosis, coagulation and fibrinolysis, pp.39-48,
 RL Plenum Press, New York (1991).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANTS SER-100; HIS-170; LEU-265 AND
 RP LYS-478.
 RA Bieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 31-491, AND CARBOHYDRATE-LINKAGE SITE.
 RA Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;
 RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some
 structural variants.";
 RL (in) Petersen H. (eds.);
 RL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,
 RL Pergamon Press, Oxford (1980).
 RN [6]
 RP SEQUENCE OF 31-491.
 RX MEDLINE=79124640; PubMed=420779;
 RA Watt K.W.K., Takagi T., Doolittle R.F.;

RT "Amino acid sequence of the beta chain of human fibrinogen.";
 RL Biochemistry 18:69-76(1979).
 RN [7]
 RP SEQUENCE OF 31-148, AND DISULFIDE BONDS.
 RX MEDLINE=76225080; PubMed=936108;
 RA Blomback B., Hessel B., Hogg D.;
 RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
 RL Thromb. Res. 8:639-658(1976).
 RN [8]
 RP SEQUENCE OF 1-38 FROM N.A.
 RX MEDLINE=87146483; PubMed=3029722;
 RA Huber P., Dalmon J., Courtois G., Laurent M., Assouline Z.,
 RA Marguerie G.;
 RT "Characterization of the 5'-flanking region for the human fibrinogen
 beta gene.";
 RL Nucleic Acids Res. 15:1615-1625(1987).
 RN [9]
 RP SEQUENCE OF 31-44.
 RA Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;
 RT "Studies on fibrinopeptides from primates.";
 RL Acta Chem. Scand. 19:1788-1789(1965).
 RN [10]
 RP REVIEW, AND DISULFIDE BONDS.
 RX MEDLINE=83254370; PubMed=6575689;
 RA Henschen A., Lottspeich F., Kehl M., Southan C.;
 RT "Covalent structure of fibrinogen.";
 RL Ann. N.Y. Acad. Sci. 408:28-43(1983).
 RN [11]
 RP DISULFIDE BONDS.
 RX MEDLINE=77245999; PubMed=891553;
 RA Gaardlund B., Hessel B., Marguerie G., Murano G., Blomback B.;
 RT "Primary structure of human fibrinogen. Characterization of
 disulfide-containing cyanogen-bromide fragments.";
 RL Eur. J. Biochem. 77:595-610(1977).
 RN [12]
 RP DISULFIDE BONDS.
 RA Doolittle R.F., Takagi T., Watt K.W.K., Bouma H. III, Cottrell B.A.,
 RA Casman K.G., Goldbaum D.M., Doolittle L.R., Friesner S.J.;
 RT "The structures of fibrinogen and fibrin.";
 RL (in) Magnusson S., Ottesen M., Foltmann B., Dano K.,
 RL Neurath H. (eds.);
 RL Regulatory proteolytic enzymes and their inhibitors, pp.163-172,
 RL Pergamon Press, New York (1978).
 RN [13]
 RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
 RX MEDLINE=84305751; PubMed=6383194;
 RA Doolittle R.F.;
 RT "Fibrinogen and fibrin.";
 RL Annu. Rev. Biochem. 53:195-229(1984).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 164-491.
 RX MEDLINE=97472408; PubMed=9333233;
 RA Spraggon G., Everse S.J., Doolittle R.F.;
 RT "Crystal structures of fragment D from human fibrinogen and its
 crosslinked counterpart from fibrin.";
 RL Nature 389:455-462(1997).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 164-491.
 RX MEDLINE=98292395; PubMed=9628725;
 RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
 RT "Crystal structure of fragment double-D from human fibrin with two
 different bound ligands.";
 RL Biochemistry 37:8637-8642(1998).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=99175089; PubMed=10074346;
 RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
 RT "Conformational changes in fragments D and double-D from human
 fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide.";
 RL Biochemistry 38:2941-2946(1999).
 RN [17]
 RP INTERACTION WITH FELN1.
 RX MEDLINE=95370284; PubMed=7642629;

RA Tran H., Tanaka A., Litvinovich S.V., Medved L.V., Haudenschield C.C.,
RA Argaves W.S.;
RT "The interaction of fibrin-1 with fibrinogen. A potential role in
RT hemostasis and thrombosis.";
RL J. Biol. Chem. 270:19458-19464 (1995).
RN [18]
RP VARIANT BALTIMORE-2 LYS-478
RX MEDLINE=89058942; PubMed=3194892;
RA Schmelzer C.H., Ebert R.F., Bell W.R.;
RT "A polymorphism at B beta 448 of fibrinogen identified during
RT structural studies of fibrinogen Baltimore II.";
RL Thromb. Res. 52:173-177 (1988).
RN [19]
RP VARIANT ISE ARG-45
RX MEDLINE=91208409; PubMed=2018836;
RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
RA Asakura S., Shirakawa S.;
RT "A new congenital abnormal fibrinogen Ise characterized by the
RT replacement of B beta glycine-15 by cysteine.";
RL Blood 77:1958-1963 (1991).
RN [20]
RP VARIANT NAPLES THR-98.
RX MEDLINE=92340664; PubMed=1634610;
RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
RT "Molecular basis of fibrinogen Naples associated with defective
RT thrombin binding and thrombophilia. Homozygous substitution of B beta
RT 68 Ala-->Thr.";
RL J. Clin. Invest. 90:238-244 (1992).
RN [21]
RP VARIANTS ILMUIDEN CYS-44 AND NIJMEGEN CYS-74.
RX MEDLINE=92228809; PubMed=1565641;
RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
RA Kerst A.F.J.A., Lord S.T.;
RT "Abnormal fibrinogens ILMUIDEN (B beta Arg14-->Cys) and NIJMEGEN (B
RT beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
RT complexes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482 (1992).
RN [22]
RP VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
RX MEDLINE=85157605; PubMed=3156856;
RA Liu C.Y., Koehn J.A., Morgan F.J.;
RT "Characterization of fibrinogen New York 1. A dysfunctional
RT fibrinogen with a deletion of B beta (9-72) corresponding exactly to
RT exon 2 of the gene.";
RL J. Biol. Chem. 260:4390-4396 (1985).
RN [23]
RP VARIANTS GLU-2; LEU-265 AND LYS-478.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [24]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373 (1999).
RN [25]
RP VARIANTS CONGENITAL AFIBRINOGENEMIA ARG-393 AND ASP-430.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonic T.,
RA Malcovati M., Mannucci P.M., Turchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause
RT congenital afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341 (2000).
RN [26]
RP VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
RX MEDLINE=21361164; PubMed=11468164;

RA Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,
RA Hangan R.R., Lord S.T.;
RT "The impaired polymerization of fibrinogen Longmont
RT [18]
RN [19]
RP VARIANT BALTIMORE-2 LYS-478
RX MEDLINE=89058942; PubMed=3194892;
RA Schmelzer C.H., Ebert R.F., Bell W.R.;
RT "A polymorphism at B beta 448 of fibrinogen identified during
RT structural studies of fibrinogen Baltimore II.";
RL Thromb. Res. 52:173-177 (1988).
RN [20]
RP VARIANT ISE ARG-45
RX MEDLINE=91208409; PubMed=2018836;
RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
RA Asakura S., Shirakawa S.;
RT "A new congenital abnormal fibrinogen Ise characterized by the
RT replacement of B beta glycine-15 by cysteine.";
RL Blood 77:1958-1963 (1991).
RN [21]
RP VARIANT NAPLES THR-98.
RX MEDLINE=92340664; PubMed=1634610;
RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
RT "Molecular basis of fibrinogen Naples associated with defective
RT thrombin binding and thrombophilia. Homozygous substitution of B beta
RT 68 Ala-->Thr.";
RL J. Clin. Invest. 90:238-244 (1992).
RN [22]
RP VARIANTS ILMUIDEN CYS-44 AND NIJMEGEN CYS-74.
RX MEDLINE=92228809; PubMed=1565641;
RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
RA Kerst A.F.J.A., Lord S.T.;
RT "Abnormal fibrinogens ILMUIDEN (B beta Arg14-->Cys) and NIJMEGEN (B
RT beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
RT complexes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482 (1992).
RN [23]
RP VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
RX MEDLINE=85157605; PubMed=3156856;
RA Liu C.Y., Koehn J.A., Morgan F.J.;
RT "Characterization of fibrinogen New York 1. A dysfunctional
RT fibrinogen with a deletion of B beta (9-72) corresponding exactly to
RT exon 2 of the gene.";
RL J. Biol. Chem. 260:4390-4396 (1985).
RN [24]
RP VARIANTS GLU-2; LEU-265 AND LYS-478.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [25]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373 (1999).
RN [26]
RP VARIANT CONGENITAL AFIBRINOGENEMIA ARG-393 AND ASP-430.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonic T.,
RA Malcovati M., Mannucci P.M., Turchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause
RT congenital afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341 (2000).
RN [27]
RP VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
RX MEDLINE=21361164; PubMed=11468164;

RA Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,
RA Hangan R.R., Lord S.T.;
RT "The impaired polymerization of fibrinogen Longmont
RT [18]
RN [19]
RP VARIANT BALTIMORE-2 LYS-478
RX MEDLINE=89058942; PubMed=3194892;
RA Schmelzer C.H., Ebert R.F., Bell W.R.;
RT "A polymorphism at B beta 448 of fibrinogen identified during
RT structural studies of fibrinogen Baltimore II.";
RL Thromb. Res. 52:173-177 (1988).
RN [20]
RP VARIANT ISE ARG-45
RX MEDLINE=91208409; PubMed=2018836;
RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
RA Asakura S., Shirakawa S.;
RT "A new congenital abnormal fibrinogen Ise characterized by the
RT replacement of B beta glycine-15 by cysteine.";
RL Blood 77:1958-1963 (1991).
RN [21]
RP VARIANT NAPLES THR-98.
RX MEDLINE=92340664; PubMed=1634610;
RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
RT "Molecular basis of fibrinogen Naples associated with defective
RT thrombin binding and thrombophilia. Homozygous substitution of B beta
RT 68 Ala-->Thr.";
RL J. Clin. Invest. 90:238-244 (1992).
RN [22]
RP VARIANTS ILMUIDEN CYS-44 AND NIJMEGEN CYS-74.
RX MEDLINE=92228809; PubMed=1565641;
RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
RA Kerst A.F.J.A., Lord S.T.;
RT "Abnormal fibrinogens ILMUIDEN (B beta Arg14-->Cys) and NIJMEGEN (B
RT beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
RT complexes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482 (1992).
RN [23]
RP VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
RX MEDLINE=85157605; PubMed=3156856;
RA Liu C.Y., Koehn J.A., Morgan F.J.;
RT "Characterization of fibrinogen New York 1. A dysfunctional
RT fibrinogen with a deletion of B beta (9-72) corresponding exactly to
RT exon 2 of the gene.";
RL J. Biol. Chem. 260:4390-4396 (1985).
RN [24]
RP VARIANTS GLU-2; LEU-265 AND LYS-478.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [25]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373 (1999).
RN [26]
RP VARIANT CONGENITAL AFIBRINOGENEMIA ARG-393 AND ASP-430.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonic T.,
RA Malcovati M., Mannucci P.M., Turchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause
RT congenital afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341 (2000).
RN [27]
RP VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
RX MEDLINE=21361164; PubMed=11468164;

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF062566; AAC16484.1; -.
DR EMBL; S79832; AAB35321.1; -.
DR EMBL; X60136; CAA42721.1; -.
DR HSSP; P08047; 1SP1.
DR MGD; MGI-198372; SP1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0003700; F:transcription factor activity; IMP.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IMP.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SMO0355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein;
KW Alternative splicing.
FT ZN_FING 624 648 C2H2-TYPE 1.
FT ZN_FING 654 678 C2H2-TYPE 2.
FT ZN_FING 684 706 C2H2-TYPE 3.
FT VARSPPLIC 57 370 Missing (in isoform 2).
FT CONFLICT 459 459 /FTId=VSP_007376.
FT SEQUENCE 781 AA; 80486 MW; 14CD12BBC58CF921 CRC64;
Alignment Scores:
Pred. No.: 4.58 Length: 781
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.85% Indels: 0
DB: 1 Gaps: 0
US-09-270-437D-8 (1-3283) x SP1_MOUSE (1-781)
QY 893 ATCTCTCGCTAGTTGGTCTCATCT 867
DB 304 ILeSer:SerAlaSerLeuValSerSer 312
RESULT 4
SP1_HUMAN STANDARD; PRT; 785 AA.
AC P08047; Q9H3Q5; Q9NR51; Q9NY21; Q9NYE7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor Sp1.
GN SP1 OR ISFP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 4-785 FROM N.A.
RC TISSUE=cervical carcinoma;
RA Haggart M.H., Ladurner A.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-558 FROM N.A.
RX MEDLINE=20545561; PubMed=10973950;
RA Takahara T., Kanazu S., Yanagisawa S., Akanuma H.;
RT "heterogeneous Sp1 mRNAs in human HepG2 cells include a product of
RL hemotypic trans-splicing.";
RL J. Biol. Chem. 275:38067-38072 (2000).
RN [3]
RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.
RX MEDLINE=88080466; PubMed=319186;
RA Kadanaga J.T., Carner K.R., Maslarsz F.R., Tjian R.;
RT "Isolation of cDNA encoding transcription factor Sp1 and functional
```

```
RT analysis of the DNA binding domain.";
RL Cell 51:1079-1090(1987).
RN [4]
RP SEQUENCE OF 1-109 FROM N.A.
RA Nicolas M., Noe V., Ciudad C.J.;
RT "Expression of transcription factor Sp1 mRNA in mammalian cells.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-98 FROM N.A.
RA Handschug K., Huebner A.;
RT "Sequencing of the 5' end of human transcription factor Sp1 mRNA.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP O-GLYCOSYLATION.
RX MEDLINE=89003041; PubMed=3139301;
RA Jackson S.P., Tjian R.;
RT "O-glycosylation of eukaryotic transcription factors: implications
for mechanisms of transcriptional regulation.";
RL Cell 55:125-133(1988).
RN [7]
RP STRUCTURE BY NMR OF 654-684 AND 684-712.
RX MEDLINE=97218212; PubMed=9065444;
RA Narayan V.A., Kziwacki R.W., Caradonna J.P.;
RT "Structures of zinc finger domains from transcription factor Sp1.
Insights into sequence-specific protein-DNA recognition.";
RL J. Biol. Chem. 272:7801-7809(1997).
RN [8]
RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
RX MEDLINE=96224025; PubMed=8626793;
RA Parks C.L., Shenk T.;
RT "The serotonin 1a receptor gene contains a TATA-less promoter that
responds to MAZ and Sp1.";
RL J. Biol. Chem. 271:4417-4430(1996).
CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
ACTIVATES MENA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
SEROTONIN RECEPTOR PROMOTER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains.
CC -!- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
PROTEINS.
CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF252284; AAP67726.1; -.
DR EMBL; AB039286; BAB13476.1; -.
DR EMBL; J03133; AAA61154.1; -.
DR EMBL; AF255682; AAF78781.1; -.
DR EMBL; AJ272134; CAB75345.1; -.
DR PIR; A29635; A29635.
DR PDB; 1SP1; 2I-APR-97.
DR PDB; 1SP2; 2I-APR-97.
DR TRANSFAC; T00759; -.
DR GlycoSuiteDB; P08047; -.
DR Genew; HGNC:11205; SP1.
DR MIM; 189906; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti...; NAS.
DR GO; GO:0016563; F:transcriptional activator activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SMO0355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
```


KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
 FT ZN_FING 626 650 C2H2-TYPE 1.
 FT ZN_FING 656 680 C2H2-TYPE 2.
 FT ZN_FING 686 708 C2H2-TYPE 3.
 FT CONFLICT 366 366 D -> G (IN REF. 3; AA SEQUENCE).
 FT CONFLICT 670 670 S -> F (IN REF. 3; AA SEQUENCE).
 FT STRAND 657 657
 FT TURN 661 662
 FT STRAND 666 666
 FT HELIX 670 677
 FT TURN 678 680
 SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;

Alignment Scores:
 Pred. No.: 4.58 Length: 785
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.85% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x SP1_HUMAN (1-785)

QY 893 ATCTTCGGCTAGTTGGTCTCATCT 867
 |||||
 Db 305 ILeSerSerAlaSerLeuValSerSer 313

RESULT 5

SP1_RAT STANDARD; PRT; 788 AA.
 ID SP1_RAT
 AC Q01714;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Transcription factor Sp1.
 GN SP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93010958; PubMed=1356762;
 RA Inatoka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
 RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
 RT "Two regulatory proteins that bind to the basic transcription element
 (BTE), a GC box sequence in the promoter region of the rat P-4501A1
 gene.";
 RL EMO J. 11:363-367(1992).
 CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
 CC RECOGNITION SITES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: O-glycosylated; contains N-acetylglucosamine side chains (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; D12768; BAA02235.1; -;
 DR PIR; JS0747; JS0747.
 DR HSP; P08047; ISPI.
 DR TRANSFAC; T00754; -;

DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF00096; zf-C2H2; 3.
 DR ProDom: PD000003; Znf_C2H2; 2.
 DR SMART: SM00355; Znf_C2H2; 3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.
 FT ZN_FING 629 653 C2H2-TYPE 1.
 FT ZN_FING 659 683 C2H2-TYPE 2.
 FT ZN_FING 689 711 C2H2-TYPE 3.
 SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Alignment Scores:
 Pred. No.: 4.58 Length: 788
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.85% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x SP1_RAT (1-788)

QY 893 ATCTTCGGCTAGTTGGTCTCATCT 867
 |||||
 Db 308 ILeSerSerAlaSerLeuValSerSer 316

RESULT 6

KROS_HUMAN STANDARD; PRT; 2347 AA.
 ID KROS_HUMAN
 AC P08922; Q15368;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase ROS precursor (EC 2.7.1.112)
 DE (c-ros-1).
 GN ROS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90280463; PubMed=2352949;
 RA Birnmeier C., O'Neill K., Riggs M., Wigler M.;
 RT "Characterization of ROS1 cDNA from a human glioblastoma cell line.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).
 RN [2]
 RP SEQUENCE OF 1790-2259 FROM N.A.
 RX MEDLINE=87064611; PubMed=3023956;
 RA Matsushima H., Wang L.-H., Shibuya M.;
 RT "Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma
 RT virus encodes for a transmembrane receptorlike molecule.";
 RL Mol. Cell. Biol. 6:3000-3004(1986).
 RN [3]
 RP SEQUENCE OF 1854-2245 FROM N.A.
 RX MEDLINE=87064625; PubMed=3785223;
 RA Birnmeier C., Birnbaum D., Walches G., Pasano O., Wigler M.;
 RT "Characterization of an activated human ros gene.";
 RL Mol. Cell. Biol. 6:3109-3116(1986).
 CC -1- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION
 CC FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
 CC receptor subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; D12768; BAA02235.1; -;
 DR PIR; JS0747; JS0747.
 DR HSP; P08047; ISPI.
 DR TRANSFAC; T00754; -;

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 0.74%
 DB: 1
 Indels: 0
 Gaps: 0

Query Match: 0.74%
 DB: 1
 Indels: 0
 Gaps: 0

US-09-270-437D-8 (1-3283) x 5H1B_PIG (1-150)

US-09-270-437D-8 (1-3283) x 5H1B_CANFA (1-161)

Qy 2418 TCTGGATCACCCTGTGTATGCAAC 2441

Qy 2418 TCTGGATCACCCTGTGTATGCAAC 2441

Db 49 SerGlySerProValTyValasn 56

Db 89 SerGlySerProValTyValasn 96

RESULT 10

RESULT 11

SH1B CANFA STANDARD; PRT; 161 AA.

Y175 STAAAM STANDARD; PRT; 171 AA.

AC P79250;

AC Q53719;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein SAV1875/SA1692/MW1815 (ORF1).

DE 5-hydroxytryptamine 1B receptor (5-HT1B) (Serotonin receptor) (5-HT1B) (5-HT1D subtype beta) (Fragment).

OS Staphylococcus aureus (strain Mu50 / ATCC 700699),

GN HTR1B.

OS Staphylococcus aureus (strain N315),

OS Canis familiaris (Dog).

OS Staphylococcus aureus (strain MW2), and

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RN [1]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=Beagle; and Alsatian; TISSUE=Artery;

RC STRAIN=MU50 / ATCC 700699, and N315;

RX MEDLINE=96296365; PubMed=8763409;

RX MEDLINE=21311952; PubMed=1418146;

RA Sgard F., Faure C., Graham D.;

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RT "Evidence for 5-HT1D alpha receptor subtype

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RL expression in canine large coronary arteries and saphenous vein.;"

RA Sekimizu K., Hirakawa H., Kuhara S., Sawano T., Inoue R.-I., Kaito C.,

CC -1- FUNCTION: This is one of the several different receptors for 5-

RA Kanehisa M., Ogasawara N., Oshima K., Goto S., Yabuzaki J.,

CC hydroxytryptamine (serotonin), a biogenic hormone that functions

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

CC as a neurotransmitter, a hormone, and a mitogen. The activity of

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

CC this receptor is mediated by G proteins that inhibit adenylyate

RT aureus.;"

CC cyclase activity.

RL Lancet 357:1225-1240 (2001).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

RN [2]

CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

RP SEQUENCE FROM N.A.

CC Strongest to the other SHT-1 subtype receptors.

RC STRAIN=MW2;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RA MEDLINE=22040717; PubMed=12044378;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

CC the European Bioinformatics Institute. There are no restrictions on its

RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,

CC use by non-profit institutions as long as its content is in no way

RA Yamamoto K., Hiramatsu K.;

CC modified and this statement is not removed. Usage by and for commercial

RT "Genome and virulence determinants of high virulence community-

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

RT acquired MRSA.;"

CC or send an email to license@isb-sib.ch)

RL Lancet 359:1819-1827 (2002).

CC [3]

RN [3]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RC STRAIN=NCTC 8325;

EX MEDLINE=94131293; PubMed=7905453;

EX MEDLINE=94131293; PubMed=7905453;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

RT "Sequence analysis of the region downstream from a peptidoglycan

CC Gene 137:253-258 (1993).

CC Gene 137:253-258 (1993).

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch)

CC or send an email to license@isb-sib.ch)

CC [3]

CC [3]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RC STRAIN=NCTC 8325;

EX MEDLINE=94131293; PubMed=7905453;

EX MEDLINE=94131293; PubMed=7905453;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

RT "Sequence analysis of the region downstream from a peptidoglycan

CC Gene 137:253-258 (1993).

CC Gene 137:253-258 (1993).

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch)

CC or send an email to license@isb-sib.ch)

CC [3]

CC [3]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RC STRAIN=NCTC 8325;

EX MEDLINE=94131293; PubMed=7905453;

EX MEDLINE=94131293; PubMed=7905453;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

RT "Sequence analysis of the region downstream from a peptidoglycan

CC Gene 137:253-258 (1993).

CC Gene 137:253-258 (1993).

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch)

CC or send an email to license@isb-sib.ch)

CC [3]

CC [3]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RC STRAIN=NCTC 8325;

EX MEDLINE=94131293; PubMed=7905453;

EX MEDLINE=94131293; PubMed=7905453;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

RT "Sequence analysis of the region downstream from a peptidoglycan

CC Gene 137:253-258 (1993).

CC Gene 137:253-258 (1993).

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch)

CC or send an email to license@isb-sib.ch)

CC [3]

CC [3]

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RC STRAIN=NCTC 8325;

EX MEDLINE=94131293; PubMed=7905453;

EX MEDLINE=94131293; PubMed=7905453;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RA Bocharde S.A., Babwan A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

RT "Sequence analysis of the region downstream from a peptidoglycan

CC Gene 137:253-258 (1993).

CC Gene 137:253-258 (1993).

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC -1- SIMILARITY: BELONGS TO THE TH13 / PFPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch)

CC or send an email to license@isb-sib.ch)

CC [3]

CC [3]

RP SEQUENCE FROM N.A.

```

DR InterPro; IPR002818; ThiJ/Pfpi.
DR Pfam; PF01965; DJ-1_Pfpi; 1.
DR TIGRFAMs; TIGR01382; Pfpi; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 18632 MW; A571A08B0FD0D719 CRC64;

Alignment Scores:
Pred. No.: 54.8 Length: 171
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.74% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x Y175_STAAM (1-171)
QY 2398 GGAAGGACACTCAGCGAGTTCTG 2421
DB 119 GlyArgThrLeuThraValLeu 126

RESULT 12
RS7_NEUCR STANDARD; PRT; 202 AA.
ID RS7_NEUCR
AC O43105;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 40S ribosomal protein S7.
GN RPS-7 OR CRPS-7.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
XN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=74-OR23-IVA;
RX MEDLINE=97174111; PubMed=9021131;
RA Vlerula P.J.;
RT "Cloning and characterization of a Neurospora crassa ribosomal
RT protein gene, crps-7";
RL Curr. Genet. 31:139-143(1997).
CC -!- SIMILARITY: Belongs to the S7E family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U73847; AAB94301.1; -
CC PIR; T46586;
CC InterPro; IPR000554; Ribosomal_S7E.
CC Pfam; PF01251; Ribosomal_S7e; 1.
CC ProDom; PD06276; Ribosomal_S7E; 1.
CC PROSITE; PS00948; RIBOSOMAL_S7E; 1.
KW Ribosomal protein.
SQ SEQUENCE 202 AA; 22819 MW; 94D87756FFB37798 CRC64;

Alignment Scores:
Pred. No.: 53.9 Length: 202
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.74% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x RS7_NEUCR (1-202)
QY 1688 AGCAGCAGGACGAGAAATACCCCTC 1711
DB 111 SerSerArgSerArgAsnThrLeu 118

RESULT 13
YSX2_CABEL STANDARD; PRT; 208 AA.
ID YSX2_CABEL
AC Q10021; Q95ZN0; Q95ZN1; Q95ZN2;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical RNA-binding protein T28D9.2 in chromosome II.
T28D9.2.
GN Caenorhabditis elegans.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
XN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=a;
CC IsoId=Q10021-1; Sequence=Displayed;
CC Name=b;
CC IsoId=Q10021-2; Sequence=VSP_005908, VSP_005910;
CC Name=c;
CC IsoId=Q10021-3; Sequence=VSP_005906, VSP_005907;
CC Name=d;
CC IsoId=Q10021-4; Sequence=VSP_005909;
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28738; AAK68314.2; -
CC EMBL; U28738; AAK68399.2; -
CC EMBL; U28738; AAK68400.1; -
CC EMBL; U28738; AAK72066.2; -
CC WormPep; T28D9.2a; CE29826.
CC WormPep; T28D9.2b; CE29827.
CC WormPep; T28D9.2c; CE28094.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rsm; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RNP_RNP_1; FALSE_NEG.
KW Hypothetical protein; RNA-binding; Alternative splicing.
FT DOMAIN 2 74
FT DOMAIN 84 205
FT DOMAIN 128
FT VARSPLIC 1 128
FT
FT MPRYLGIKIPYNGHGERDVERFLKGYGKINNISMRYGFAPVD
FT PEDSRDAEDACHLDLDTGSGSMRLVVMARGLKPRGNDRH
FT GSRSPRRSRSPRRSRTPPRSRSRDRKSRSRSS
FT RSRSP -> MLKTHAMIMKPEKALCALSKWLAENHAA
FT MTVMRDHDDVDVALLDVLVHQLQEDVALLVIANVLADPVP
FT EAPPDPDHFQVVAENPVRARRKTELWVNIQYIISIFSP
FT SPHSYQKELFLKE (in isoform c).
FT /FTId=VSP_005906.
FT Missing (in isoform c).
FT /FTId=VSP_005907.
FT LKRESRSRSLPAXDRSRT -> FHSSMRNQYHLQAHIAI
FT AYT (in isoform b).
FT

```

```

FT VARSPLIC 148 208 /FTID=VSP_005908.
KRAASRSRSLPAKDRSRTRSGPPKNGGDKRSVSRGRSH
SRGNSVSRSRSPGSPKD -> OLANISVSYSPLDSD
SIRKNVILSVKRDVLVHCQOKIAPPEVDLPQRTAIE
SAA (in isoform d).
/FTID=VSP_005909.
Missing (in isoform b).
/FTID=VSP_005910.
SQ SEQUENCE 208 AA; 23946 MW; A33CAFC78D74A6F8 CRC64;

Alignment Scores:
Pred. No.: 53.8 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x YSX2_CAREL (1-208)

QY 202 AGTCACGAGGGCTAGCCGAGCT 179
|||||
DB 93 SerProArgArgSerArgThr 100

RESULT 14
RS2_METMA STANDARD; PRT; 224 AA.
AC Q8PFW41;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MM1760.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierer A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -1- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB013412; AM31456.1; -
DR HAMAP; MF_00291; -; 1.
DR InterPro; IPR001865; Ribosomal_S2.
DR InterPro; IPR005707; Ribosomal_S2_e/a.
DR Pfam; PF00318; Ribosomal_S2_1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRfams; TIGR01012; Sa_S2_E_A; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 224 AA; 24625 MW; 144D106450EA125E CRC64;

Alignment Scores:
Pred. No.: 53.4 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 1 Gaps: 0

```

```

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-8 (1-3283) x RS2_METMA (1-224)

QY 2367 CGGTGGCTCAAGCTCTCTCT 2344
|||||
DB 83 ArgValAlaSerLysLeuLeuSer 90

RESULT 15
RS2_METAC STANDARD; PRT; 225 AA.
AC Q8TT39;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MA0500.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atcoor D., Brown A.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity."
RL Genome Res. 12:532-542(2002).
CC -1- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB010721; AA04044.1; ALT_INIT.
DR HAMAP; MF_00291; -; 1.
DR InterPro; IPR001865; Ribosomal_S2.
DR InterPro; IPR005707; Ribosomal_S2_e/a.
DR Pfam; PF00318; Ribosomal_S2_1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRfams; TIGR01012; Sa_S2_E_A; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 225 AA; 24779 MW; 02C722D26A5F411F CRC64;

Alignment Scores:
Pred. No.: 53.4 Length: 225
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.76% Indels: 0
DB: 1 Gaps: 0

```

US-09-270-437D-8 (1-3283) x RS2_METAC (1-225)

QY 2367 CGGTGGGCTCAAAGCTCTCTCT 2344

Db |||||||||||||||||||
84 ArgValAlaSerLysLeuLeuSer 91

Search completed: August 6, 2004, 13:22:24
Job time : 42.8809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 6, 2004, 13:12:26 ; Search time 156.647 Seconds
(without alignments)
13225.208 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 1077
Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaatgtttattt 3283

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1
Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPT.spool_p/US09270437/runat_06082004_141220_1815/app_query.fasta_1.5582
-DB=SPREMBL 25 -QFWT=fastan -SUFFIX=oligo806.rspt -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437@cgn_1_1_287@runat_06082004_141220_1815 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	556	51.6	556 4 Q9Y6M1	Q9Y6M1 homo sapien

2	88	8.2	545	11	Q7TQF9
3	23	2.1	582	13	Q9PW80
4	23	2.1	593	13	O57526
5	23	2.1	594	13	O73932
6	19	1.8	576	13	O42254
7	19	1.8	577	4	Q9NZ18
8	19	1.8	577	11	Q8CGX0
9	19	1.8	577	11	Q8BRH1
10	19	1.8	577	11	Q80US9
11	19	1.8	577	11	O88477
12	19	1.8	579	4	O00425
13	19	1.8	579	11	Q8C2J9
14	19	1.8	579	11	Q9CPN8
15	14	1.3	169	11	Q7TP50
16	12	1.1	100	4	Q86VB1
17	10	0.9	215	16	Q824H3
18	10	0.9	1328	4	Q9UW1
19	9	0.8	319	16	Q9RV01
20	9	0.8	329	3	Q8NIY1
21	9	0.8	366	16	Q926X1
22	9	0.8	385	11	Q9DAE2
23	9	0.8	422	16	O8FLQ8
24	9	0.8	451	5	Q18705
25	9	0.8	512	16	Q8P646
26	9	0.8	549	10	Q94D52
27	9	0.8	559	11	Q8K4R0
28	9	0.8	623	10	Q9CAY1
29	9	0.8	704	5	Q9V113
30	9	0.8	711	17	Q8TVS7
31	9	0.8	784	11	Q89087
32	9	0.8	798	4	Q86TN8
33	9	0.8	807	10	Q9U672
34	9	0.8	893	5	Q81429
35	9	0.8	1024	10	Q9XGS7
36	9	0.8	1086	16	Q8EPB9
37	9	0.8	1135	10	Q7XS11
38	9	0.8	1226	10	Q9LXT8
39	9	0.8	1854	11	Q63131
40	9	0.8	2317	11	Q63130
41	9	0.8	2338	11	Q63132
42	9	0.8	8625	5	Q86GD6
43	8	0.8	37	4	Q8TD81
44	8	0.8	37	8	O78854
45	8	0.8	45	2	Q8GF59

ALIGNMENTS

RESULT 1

Q9Y6M1	PRELIMINARY;	PRT;	556 AA.
AC	Q9Y6M1;		
DT	01-NOV-1999 (TReMBLrel. 12, Created)		
DT	01-NOV-1999 (TReMBLrel. 12, Last sequence update)		
DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)		
DE	Hepatocellular carcinoma autoantigen.		
GN	P62.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
[1]	SEQUENCE FROM N.A.		
RP	MEDLINE=99207072; PubMed=10190901;		
RX	Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;		
RA	"A novel cytoplasmic protein with RNA-binding motifs is an autoantigen		
RT	in human hepatocellular carcinoma."		
RL	J. Exp. Med. 189:1101-1110(1999).		
CC	-1- SIMILARITY: CONTAINS 4 KH DOMAINS.		
CC	EMBL; AF057352; AAD31596.1; "		
DR	GO; GO:0005737; Cytoplasm; TAS.		
DR	GO; GO:0003723; F:RNA binding; TAS.		
DR	GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.		


```

RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., McEwan P.J., McKernan K.J., Abramson R.D., Mullany S.J.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054552; AAH54552.1; -.
KW Hypothetical protein.
SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;

Alignment Scores:
Pred. No.: 5,66e-83 Length: 545
Score: 88.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.17% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-8 (1-3283) x Q7TQF9 (1-545)
QY 628 TCTCAGGCCGACAGATGTGATTTCCCGCTCGGATCTGTCCGCCACCCAGTTGTGTGT 687
Db 179 SerGlnAlaArgGlnIleAspPheProLeuArgIleuValProThrGlnPheValGly 198
QY 688 GCCATCATCGGAAGAGGGCTTGACCATAAAGACATCTAAGCAGACCCAGTCCCGG 747
Db 199 AlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnSerArg 218
QY 748 GTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAGCTCTCACCATCCATGCC 807
Db 219 ValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleHisAla 238
QY 808 ACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAATCATGACAGAAAGGGA 867
Db 239 ThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysGluAla 258
QY 868 GATGAGACCAAACTAGCCGAGAG 891
Db 259 AspGluThrLysLeuAlaGluGlu 266

RESULT 3
Q9PW80 PRELIMINARY; PRT; 582 AA.
ID AC Q9PW80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vg1 RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding
DE protein).
GN DVR1RBP.

```

```

OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955.
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
RA Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;
RT "Vg1 RBP intracellular distribution and evolutionarily conserved
RT expression suggest multiple roles during development.";
RL Mech. Dev. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., McEwan P.J., McKernan K.J., Abramson R.D., Mullany S.J.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF161270; AAD45610.1; -.
DR EMBL; BC045873; AAH45873.1; -.
DR ZFIN; ZDB-GENE-000308-1; dvrlrbp.
DR GO; GO:0003676; F:nucleic acid binding; IBA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 582 AA; 63351 MW; 9DAB63200681B306 CRC64;

Alignment Scores:
Pred. No.: 2,72e-14 Length: 582
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.14% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-8 (1-3283) x Q9PW80 (1-582)
QY 1378 GGGCCACCGGAAGCCAGTTCAGGCGCCAGGCGCGGATCTTTGGGAACTGAAAGAGAA 1437
Db 461 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 480
QY 1438 AACTCTTTT 1446
|||||

```

```

Db      481 AsnPhePhe 483
RESULT 4
O57526
ID      O57526      PRELIMINARY;      PRT;      593 AA.
AC      O57526;
DT      01-JUN-1998 (TrEMBLrel. 06, Created)
DT      01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      KH domain-containing transcription factor B3.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92249652; PubMed=1577195;
RA      Pfaff S.L., Taylor W.L.;
RT      "Characterization of a Xenopus oocyte factor that binds to a
RT      developmentally regulated cis-element in the TFIID gene.";
RL      Dev. Biol. 151:306-316(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Griffin D., Taylor W.L.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA      Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL      Genes Dev. 0:0-0(1998).
CC      -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR      EMBL; AF042353; AAB97457.1; -.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      InterPro; IPR004087; KH_dom.
DR      InterPro; IPR004088; KH_type_1.
DR      Pfam; PF00013; KH; 4.
DR      Pfam; PF00076; rrm; 2.
DR      SMART; SM00322; KH; 2.
DR      SMART; SM00360; RRM; 2.
DR      PROSITE; PS00084; KH_TYPE_1; 4.
DR      PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ      SEQUENCE 593 AA; 65385 MW; 5A5AB4B4A1D55DF7 CRC64;

Alignment Scores:
Pred. No.:      2,71e-14      Length:      593
Score:          23.00      Matches:      23
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    2.14%      Indels:      0
DB:             13      Gaps:        0

US-09-270-437D-8 (1-3283) x O57526 (1-593)
QY      1501 GCTGCGCGGTGATTGGCAAGCTGGCAAGACCGTGACGAACTGCAGAACTTAACCACT 1560
Db      511 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 530
QY      1561 GCAGAGATC 1569
Db      531 AlaGluVal 533

RESULT 5
O73932
ID      O73932      PRELIMINARY;      PRT;      594 AA.
AC      O73932;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      VGI RNA binding protein variant D.

OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92249652; PubMed=1577195;
RA      Pfaff S.L., Taylor W.L.;
RT      "Characterization of a Xenopus oocyte factor that binds to a
RT      developmentally regulated cis-element in the TFIID gene.";
RL      Dev. Biol. 151:306-316(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Griffin D., Taylor W.L.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA      Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL      Genes Dev. 0:0-0(1998).
CC      -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR      EMBL; AF042353; AAB97457.1; -.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      InterPro; IPR004087; KH_dom.
DR      InterPro; IPR004088; KH_type_1.
DR      Pfam; PF00013; KH; 4.
DR      Pfam; PF00076; rrm; 2.
DR      SMART; SM00322; KH; 2.
DR      SMART; SM00360; RRM; 2.
DR      PROSITE; PS00084; KH_TYPE_1; 4.
DR      PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ      SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;

Alignment Scores:
Pred. No.:      2,71e-14      Length:      594
Score:          23.00      Matches:      23
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    2.14%      Indels:      0
DB:             13      Gaps:        0

US-09-270-437D-8 (1-3283) x O73932 (1-594)
QY      1501 GCTGCGCGGTGATTGGCAAGCTGGCAAGACCGTGACGAACTGCAGAACTTAACCACT 1560
Db      512 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 531
QY      1561 GCAGAGATC 1569
Db      532 AlaGluVal 534

RESULT 6
O42254
ID      O42254      PRELIMINARY;      PRT;      576 AA.
AC      O42254;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Zipcode-binding protein.
GN      ZBP1.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97220007; PubMed=9121465;
RA      Ross A.F., Olevnikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;
RT      "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL      Mol. Cell. Biol. 17:2158-2165(1997).
CC      -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR      EMBL; AF026527; AAB82295.1; -.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.

```

DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM RNP 1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AA2D1D81C8811 CRC64;

Alignment Scores:
Pred. No.: 4,58e-10 Length: 576
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-8 (1-3283) x 042254 (1-576)

QY 1501 GCTGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAACC 1557
Db 498 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeuThr 516

RESULT 7

Q9NZ18 PRELIMINARY; PRT; 577 AA.
ID Q9NZ18;
AC Q9NZ18;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE mRNA-binding protein CRBP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ioannidis P., Tragas T., Dimitriadis E., Samiotaki M.,
RA Panoutsakopoulos G., Kyriazoglou I., Voutoulas S., Tsipalis C.M.,
RA Kittas C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
RT homologous to both human IMP-1 and mouse CRD-BP, in benign
RT malignant mesenchymal tumors.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF198254; AAP37203.1; -.
DR HSSP; P11940; 1CVJ.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;

Alignment Scores:
Pred. No.: 4,58e-10 Length: 577
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-8 (1-3283) x Q9NZ18 (1-577)

QY 1360 AGGATGTCATCATCCAGCGGCCACCGAAGCCAGTTCAGGCCCGACGCGATC 1416

Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
RESULT 8
Q8CGX0 PRELIMINARY; PRT; 577 AA.
ID Q8CGX0;
AC Q8CGX0;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE B-actin zipcode binding protein 1.
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Eom T., Singer R.H., Bassell G.J.;
RT "Molecular interactions between rZBP1 and b-actin zipcode required for
RT transport of mRNA and stimulation of spine growth.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF541940; AA016210.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1EE CRC64;

Alignment Scores:
Pred. No.: 4,58e-10 Length: 577
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.76% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-8 (1-3283) x Q8CGX0 (1-577)

QY 1501 GCTGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAACC 1557
Db 499 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 9

Q8BRH1 PRELIMINARY; PRT; 577 AA.
ID Q8BRH1;
AC Q8BRH1;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGF2BP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK044850; EAC32119.1; -.
DR MGD; MGI:1890357; Igf2bp1.

[5] RN RP
SEQUENCE FROM N.A.
RA Rosa J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
RA Fleisig A.J. (in press)
RRL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
[6] RN RP
SEQUENCE FROM N.A.
RRC STRAIN=C57BL/6J; TISSUE=Embryo;
RRC MEDLINE=21085660; PubMed=11217851;
RRC

CC	-1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR	EMBL; U97188; AAC35208.1; -.
DR	EMBL; U76705; AAD09223.1; -.
DR	GO; GO:0005737; C:cytosolasm; TAS.
DR	GO; GO:0003723; F:rRNA binding; TAS.
DR	GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR	GO; GO:0006412; P:protein biosynthesis; TAS.
DR	GO; GO:0006396; P:rRNA processing; TAS.
DR	InterPro; IPR004087; KH_dom.
DR	InterPro; IPR004088; KH_type_1.
DR	InterPro; IPR00504; RNA_rec_mot.
DR	Pfam; PF00013; KH; 4.
DR	Pfam; PF00076; rrm; 2.
DR	SMART; SM00322; KH; 4.
DR	SMART; SM00360; RRM; 2.
DR	PROSITE; PS50084; KH_TYPE_1; 4.
DR	PROSITE; PS50102; RRM; 2.
DR	PROSITE; PS50030; RNP_RNP_1; FALSE_NEG.
SQ	SEQUENCE 579 AA; 63720 MW; AB5C3A8EE3C135C5 CRC64;
 Alignment Scores:	
Pred. No.:	4,58e-10 Length: 579
Score:	19.00 Matches: 19
Percent Similarity:	100.00% Conservative: 0
Best local Similarity:	100.00% Mismatches: 0
Query Match:	1.76% Indels: 0
DB:	4 Gaps: 0
 US-09-270-437D-8 (1-3283) x O00425 (1-579)	
OY	1360 AGATGGTCATCATCACCGGCGCAACCGAGTTCAGGCCCAGGACGGATC 141
Dd	 452 ArgMetValIleIleThrGlyProProgluAlaGlnPheLysalaGlnGlyArgile 470
 RESULT 13	
QC8CJ9	PRELIMINARY; PRT; 579 AA.
ID	QC8CJ9
AC	QC8CJ9;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Insulin-like growth factor 2.
GN	IGF2BP3.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;	[1]
SEQUENCE FROM N.A.	
RC	STRAIN=NOD; TISSUE=Thymus;
RA	MEDLINE=22354663; PubMed=12466851;
RX	The FANTOM Consortium,
RA	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs."
RL	Nature 420:563-573 (2002).";
DR	EMBL; AK088465; BAC40370.1; -.
DR	MGI; MGI:1890359; IGF2bp3.
DR	GO; GO:0003676; P:nucleic acid binding; IEA.
DR	InterPro; IPR004087; KH dom.
DR	InterPro; IPR004088; KH_type_1.
DR	InterPro; IPR00504; RNA_rec_mot.
DR	Pfam; PF00013; KH; 4.
DR	Pfam; PF00076; rrm; 2.
DR	SMART; SM00322; KH; 4.
DR	SMART; SM00360; RRM; 2.
DR	PROSITE; PS50084; KH_TYPE_1; 4.
DR	PROSITE; PS50102; RRM; 2.
SQ	SEQUENCE 579 AA; 63551 MW; 937E601A95D06B77 CRC64;
 Alignment Scores:	
Pred. No.:	4,58e-10 Length: 579
Score:	19.00 Matches: 19

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.76% Indels: 0
 DB: 11 Gaps: 0

US-09-270-437D-8 (1-3283) x Q8C2J9 (1-579)

QY 1360 AGGATGTCATCATCCCGGACCGGACCGGACCGGACCGGACCGGACCGGATC 1416
 |||||
 Db 452 ArgMetValIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 14

Q9CPN8 PRELIMINARY; PRT; 579 AA.
 AC Q9CPN8
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library,
 DE clone.2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
 DE (Insulin-like growth factor 2, binding protein 3).
 GN IGFBP3 OR 2610036B18 OR MIMF3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojoberi T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Havaehizaki Y.
 RT Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
 RA Yuasa Y., Takeda M., Okano H.;
 RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications
 RT for the developing central nervous system.";
 RL J. Neurosci. Res. 0:0-0(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22338257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AK011689; BAB27779.1; -
 DR EMBL; AB046173; BAB19755.1; -
 DR EMBL; BC045138; AAK45138.1; -
 DR EMBL; BC049082; AAK49082.1; -
 DR MGD; MGI:1890359; Igf2bp3.

DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rim; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS00084; KH TYPE 1; 4.
 DR PROSITE; PS0102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Alignment Scores:

Pred. No.: 4.59e-10 Length: 579
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.76% Indels: 0
 DB: 11 Gaps: 0

US-09-270-437D-8 (1-3283) x Q9CPN8 (1-579)

QY 1360 AGGATGTCATCATCCCGGACCGGACCGGACCGGACCGGACCGGATC 1416

|||
 Db 452 ArgMetValIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 15

Q7TP50 PRELIMINARY; PRT; 169 AA.
 AC Q7TP50
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ab2-255.
 OS Rattus norvegicus [Rat].
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
 RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
 RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
 RT "Liver regeneration after PH.";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY325199; AAP2600.1; -
 SQ SEQUENCE 169 AA; 17743 MW; 19BF6295C1000CBA CRC64;

Alignment Scores:

Pred. No.: 0.000102 Length: 169
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0

DB: 11 Gaps: 0
US-09-270-437D-8 (1-3283) x Q7TP50 (1-169)
QY 1360 AGGATGGTCATCATCACCGGCGAAGCCCGTTCAG 1401
Db 61 ArgMetValIleIleThrGlyProProGluAlaGlnPheLys 74

Search completed: August 6, 2004, 13:31:04
Job time : 176.647 secs

